The Dissertation Committee for Sun Young Park certifies that this is the approved version of the following dissertation:

A STUDY ON DIAGNOSTIC IMAGE ANALYSIS FOR THE DETECTION OF PRECANCEROUS LESIONS USING MULTI-SPECTRAL DIGITAL IMAGES

Committee:

Rebecca Richards-Kortum, Supervisor

Mia K. Markey, Co-Supervisor

E. Neely Atkinson

Michele Follen

Calum MacAulay

Joydeep Ghosh

A STUDY ON DIAGNOSTIC IMAGE ANALYSIS FOR THE DETECTION OF PRECANCEROUS LESIONS USING MULTI-SPECTRAL DIGITAL IMAGES

by

Sun Young Park, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2007

Dedication

To my parents, Jong Bae Park and Young Sook Kim, and my husband, Manho Joung.

Acknowledgements

None of this work would have been possible without the selfless assistance of a great number of people. I surely need an entire chapter for acknowledging all of them. Therefore, in order for both ink and paper conservation, I would like to issue a sincere thank you to all of the professors, doctors, nurses, instructors, technicians, students, staffs, and friends who helped to bring me to where I am today.

I would also like to express my most sincere gratitude to my advisor, Dr. Rebecca Richards-Kortum for her vision, leadership, guidance, encouragement, support and great ideas throughout the whole process of my research. She created an excellent working environment and made this experience enjoyable and rewarding. Rebecca is a wonderful mentor, and her commitment to teaching and dedication to her students was an inspiration to always strive to become a better scientist. I have always admired Rebecca's dedication to establishing a tradition of women scientists. I will be forever in debt to Rebecca for all that I have learned from her.

I would like to gratefully thank the members of my committee for their valued guidance, time, helpful discussion and contribution to this work. Dr. Michele Follen and her clinical staffs in MD Anderson provided tremendous support for doing research in the clinic; Dr. Mia Markey has been a great advisor with her invaluable guidance and advice on my algorithm design. I would like to recognize Dr. Neely Atkinson, Dr.

Calum MacAulay and Dr. Joydeep Ghosh for their wonderful support and all technical advice. I am honored to have all of them as my dissertation committee members.

I would also like to thank the students, post-docs, and staff of the Optical Spectroscopy and Imaging Laboratory for their technical support and for the wonderful working environment. I am very grateful to having worked with such a wonderful and inspiring group of people.

In the end, I would like to thank my family. I am forever grateful to my parents Jong Bae and Yong Sook. It was through their guidance and examples that I became the person I am today. My special thanks go to my brother, Yong Hwan, for having visited me regardless of tremendous flight hours and for always being supportive. Last but not the least, I would like to thank my husband, Manho, for believing in me and for continuously supporting and loving me as well as encouraging me at each and every step that made this dissertation possible.

A STUDY ON DIAGNOSTIC IMAGE ANALYSIS FOR THE DETECTION OF PRECANCEROUS LESIONS USING MULTI-SPECTRAL DIGITAL IMAGES

Publication No._____________

Sun Young Park, Ph.D.
The University of Texas at Austin, 2007

Supervisor: Rebecca Richards-Kortum
Co-Supervisor: Mia K. Markey

This dissertation explores a diagnostic image analysis framework using multispectral digital colposcopy for real-time *in vivo* detection of cervical cancer. In the first part of the dissertation, the clinical feasibility of a previously developed multispectral digital colposcope (MDC) is demonstrated using a hamster cheek pouch model of carcinogenesis. Various studies on MDC applications to cervical cancer detection in human subjects are then presented. First, an automated diagnostic image analysis algorithm for cervical cancer using white light reflectance images is presented. The algorithm can identify pre-neoplastic tissue areas from an entire cervix based on intensity changes feature in the reflectance images induced by acetic acid treatment. Then, the information about tissue type is incorporated into the diagnostic image analysis framework. For this purpose, a Markov Random Field (MRF) model is adopted and the results are discussed. One of the practical difficulties of utilizing a MRF model in unpolarized white light reflectance imaging is the specular reflection problem since the effect of specular reflection extends into surrounding tissue areas. Through the use of cross polarized imaging, the effects of specular reflection reduced and the ability to segment images based on tissue types is enhanced, leading to better diagnostic performance. The diagnostic performance of polarized imaging is compared to that of unpolarized imaging. In order to assess the performance of the proposed approach, a gold standard for the entire cervical image is constructed using histopathology results from a whole cervix specimen.

The results presented in this dissertation indicate that an automated diagnostic image analysis framework for early detection of cervical cancer has the potential to be clinically applied as a low cost alternative screening technique in developing countries. Advances in imaging technology as well as in image analysis algorithms will continue to reduce the cost of diagnostic imaging systems and improve the imaging and diagnostic capability, leading to an inexpensive, real-time, minimally-invasive alternative to conventional screening techniques for early detection of cervical cancer in developing countries.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 MOTIVATION

Cancer is the second most common cause of death after heart disease worldwide. In 2007, the American Cancer Society estimated 1,444,920 new cases and 559,650 deaths in the Unites States [1]. Tremendous efforts have been devoted to cancer detection and treatment and, as a result, annual cancer deaths in developed countries including the Unites States have recently decreased. In particular, cervical cancer is one of the cancers whose death cases in developed countries has significantly decreased due to the early detection programs as, in many developed countries, screening programs for cervical cancer have been utilized successfully. However, due to the lack of resources, expensive infrastructure for cervical cancer screening in the developing world is not easily accessible, and this causes the risk of cervical cancer deaths in many developing countries to remain considerably high. Currently, about 80% of death cases from cervical cancer occur in developing countries, and the portion of cervical cancer death cases in developing countries is estimated to increase to 90% by 2020 [2-4]. This suggests that cervical cancer reflects and will reflect global health inequity significantly.

Responding to this, many researchers have focused on developing low cost screening techniques for cervical cancer which can be clinically applied in developing countries [5-24]. Simple techniques such as Direct Visual Inspection (DVI) and Visual Inspection with Acetic Acid (VIA) have been extensively studied as alternative techniques [6-15], and, recently, applications of optical technology have emerged as a new alternative [16-24]. Multispectral digital imaging is one optical technique that shows promise for the detection of cervical intraepithelial neoplasia. In particular, automated diagnostic analysis of colposcopic images can provide real-time invivo screening which can be an inexpensive and quantitative alternative to existing screening tools.

The goal of this study is to develop a diagnostic tool which can automatically identify neoplastic tissue from digital colposcopy images. Various topics including acetowhitening feature analysis, diagnostic consideration of tissue types, polarized and unpolarized white light reflectance imaging, and fluorescence imaging are considered. The images that are analyzed and used to develop the diagnostic tool were obtained using two previously developed multispectral digital colposcopes (MDCs), which acquire white light reflectance and fluorescence images of the cervix.

1.2 SPECIFIC AIMS

The main objective of this research is to develop a diagnostic image analysis framework using autofluorescence and reflectance images acquired by the MDCs for detection of epithelial pre-cancers or cancers.

Four specific aims of this research are as follows.

Aim 1: Develop diagnostic image analysis algorithms based on the intensity change feature of the reflectance images due to acetic acid treatment.

Aim 2: Develop image classification algorithms for autofluorescence images measured using the MDC and explore the applicability of the auto fluorescence images for diagnostic purposes based on the developed algorithms.

Aim 3: Apply polarized imaging for specular reflection removal and compare the diagnostic performance of polarized and unpolarized imaging.

Aim 4: Propose a pathology mapping method for a gold standard for the entire cervical image and a quantitative measure for evaluating the diagnostic performance of the developed algorithms.

1.3 DISSERTATION OVERVIEW

This dissertation is composed of 7 chapters. Chapter 1 (this chapter) is a general introduction of this research. Chapter 2 provides background on cervical cancer, optical techniques, MDC and recent literature on diagnostic image analysis tools for cervical cancer or precancer detection using multispectral imaging. In chapter 3, a study of a hamster cheek pouch model of carcinogenesis using multispectral digital images is utilized and the results are discussed. From chapter 4 to chapter 6, MDC applications to cervical cancer detection in human subjects are presented. Chapter 4 presents an automated diagnostic image analysis framework for cervical cancer detection using intensity changes in white light reflectance images. The results of the proposed approach are also discussed. Chapter 5 proposes an image analysis framework using a Markov Random Field (MRF) model and reports the diagnostic performance of the approach. A MRF framework is proposed in order to incorporate the information of tissue type into diagnosis procedures. In chapter 6, polarized imaging is applied for specular reflection removal and the diagnostic performances of polarized and unpolarized imaging are compared. Finally, chapter 7 provides a summary of results and conclusions for the research presented in this dissertation. Future directions of the proposed methods are also provided in the chapter.

Chapter 2: Background

2.1 INTRODUCTION

The goals of this study are to develop a novel diagnostic image analysis framework in order to improve detection of cervical neoplasia and to show the validity of the framework by applying the framework to clinical studies. In Chapter 2, I will address the clinical significance of cervical cancer and the potential of multispectral digital colposcopy and diagnostic image analysis algorithm. The design and specifications of the multispectral digital colposcope systems used for this dissertation will be presented. Additionally, various studies testing diagnostic image analysis algorithms will be summarized.

2.2 EARLY DETCTION OF CERVICAL CANCER

Cervical cancer ranks second to breast cancer in incidence in women worldwide and is a leading cause of cancer-related death in women in developing countries [1, 3]. Each year, approximately 500,000 cases of cervical cancer are diagnosed worldwide, and more than 270,000 women die somewhere in the world [1, 3].

Most cases of cervical cancer develop slowly. The precursors, denoted as Cervical Intraepithelial Neoplasia (CIN) or Squamous Intraepithelial Lesions (SIL), can take three to twenty years to develop into invasive cervical cancer. These precursors, which may or may not develop into cancer, can be detected through screening test such as a Papanicolaou's smear and can be treated before they develop into cancer. In developed countries, the incidence and the death rates of cervical cancer have declined over the past 50 years, mainly due to increased screening programs for cancerous or pre-cancerous lesions. The United States has reported a decrease of up to 50% in the incidence and the death rates of cervical cancer [24].

However, in developing countries, the incidence and the mortality rates are still relatively high, since cervical cancer prevention programs are rarely implemented. The absence of screening programs in many developing countries is due to the lack of resources. Many current screening and diagnostic techniques rely on sophisticated laboratory procedures performed by highly trained personnel [6-8].

Alternatively, the use of simple techniques such as Direct Visual Inspection (DVI) or Visual Inspection with Acetic Acid (VIA) have been explored in many developing countries [8, 9, 11-15]. Recent reviews of the performance of these methods have reported sufficient sensitivity and specificity when performed by trained professionals indicating they may serve as viable alternatives to Pap smear screening in low-resource settings [9, 11-13]. DVI or VIA are both techniques that allow a provider to visualize the cervix with either the naked eye or low power magnification after the application of acetic acid [5-7]. These techniques are simple and inexpensive, and do not require a laboratory infrastructure since they utilize see-and-treat strategies. However, the performance of these methods depends highly on the providers' expertise [13-15]. It is crucial to define objective criteria for the positive identification of a lesion and to train personnel to correctly implement screening programs based on these criteria. Recent reviews of the performance of these methods have reported sufficient sensitivity and specificity when performed by trained professionals and, therefore, they may serve as viable alternatives to Pap smear screening in low-resource settings [6-12].

2.3 OPTICAL TECHNIQUES

Among various organs, the cervix has been especially regarded as attractive for applications of optical detection techniques for cancer due to its easy access for imaging and due to the capability of interrogation of the changes in the epithelium and stroma using visible and near infrared light [18, 19, 21]. Therefore, optical techniques have been investigated as potential diagnostic tools for cervical cancer with possible cost benefits in many studies [18-23].

Recently, reflectance and fluorescence spectroscopy have been studied as potential diagnostic techniques for cervical cancer and its precursors. Evidence in the literature indicates that reflectance and fluorescence spectroscopy can be used to identify pre-cancerous lesions [18-23]. The advantage of reflectance and fluorescence spectroscopy over current detection methods for cervical cancer includes the ability to quantitatively detect changes in cellular chemistry and tissue architecture associated with progression of disease in a fast and nondestructive manner [18-23].

There are primarily two different approaches for reflectance and fluorescence spectroscopy: point spectroscopy and multi-spectral imaging. Point spectroscopy which uses a fiber optic probe in direct contact with the tissue, has demonstrated relatively high diagnostic performance. In diagnostic colposcopy clinics, reported sensitivity and specificity are 86% and 74%, respectively; in screening settings, they were 75% and 80%, respectively [18-21]. In a large clinical study, Huh et al. [23] measured the performance of optical detection of high-grade intraepithelial neoplasia (HGSIL) from 604 patients, finding a sensitivity of 90% and a specificity of 70%. In a randomized clinical trial, Alvarez et al. [25] investigated the impact of an optical detection system that utilizes fluorescence, white light reflectance, and video imaging. This optical detection system, used as an adjunct to colposcopy, showed an increased ability to discriminate HGSIL from normal tissue compared to colposcopy alone.

Variability in the measurement of fluorescence spectroscopy has previously been studied using point spectroscopy. Differences in histological diagnosis, menopausal status, and age among patients can cause variability in the measurements of fluorescence spectroscopy, while smoking, race, and menstrual cycle do not appear to be significant sources of variability [21-23].

In spite of its promising results, point spectroscopy has a drawback: an extremely limited field of view (less than 2mm). With this limitation, it is difficult to apply point spectroscopy to assess the entire field at risk; thus, a screening process is needed in order to identify a relatively small area at risk to be probed by point spectroscopy. Multi-spectral imaging can provide sufficient field of view to evaluate the entire surface of the cervix. Furthermore, multi-spectral imaging usually results in digital images as the measurement outcome and enables one to apply well-developed digital image processing techniques for diagnostic purposes. Therefore, many clinical applications of medical imaging systems have been reported based on multi-spectral imaging [26-40].

In multi-spectral imaging applications, there are two major components: an imaging system for measurement and image analysis algorithms. Each of these components is described in the following two subsections. In subsection 2.4, the multispectral digital colposcope systems used in this study are described as an illustration of more general imaging systems. Subsection 2.5 presents a summary of various approaches reported in literature focusing on image analysis techniques.

2.4 MULTISPECTRAL DIGITAL COLPOSCOPE

An imaging system is required for clinical applications of field spectroscopy and, as illustrations, I describe two imaging systems used for this research in this section. One is a pilot MDC developed to measure multispectral reflectance and autofluorescence images in a short period of time [39]. As a multi-spectral imaging device, it can assess the entire field at risk and detect biochemical and morphological changes as cancer develops. After investigating the feasibility of the pilot MDC system, a new MDC system was developed to expand the range of optical measurements that could be made. In this dissertation, the pilot and the new MDCs were utilized for collecting fluorescence and white light reflectance images. More detail regarding the two MDCs is provided in the following subsections.

2.4.1 Pilot MDC

The pilot MDC is a low cost device designed to be compact and transportable. The pilot MDC is composed of a light source, a standard colposcope, and a video rate color CCD camera with a frame grabber. A system diagram and photograph of the pilot MDC system are shown in Figure 2.1 (a) and (b), respectively [39].

The pilot MDC was built around a commercially available, tilt-stand colposcope (Model 1DL, Leisegang, Germany). This colposcope produces stereoscopic vision at 7.5X magnification. The original halogen lamp of the colposcope was removed to accommodate a fiber optic light guide, coupled to a Xenon arc lamp, which provides broadband illumination for reflectance imaging and monochromatic illumination for fluorescence imaging. Monochromatic light for fluorescence excitation is produced using 20 nm bandwidth band-pass filters (Chroma Technology Corp., Rockingham, VT) centered at 345 nm and 440 nm excitation enclosed in a motorized filter wheel. The optical density of each band-pass filter was at least 4 in the out of band regions. Long pass filters (Omega Optical Inc., Brattleboro, VT) at 388 nm and 468 nm block scattered excitation light during collection of fluorescence images. Each long-pass filter has an optical density of at least 5 below the cut-on wavelength. The reflectance and the fluorescent light are captured and filtered using a commercially available video rate color charge-coupled device (CCD) camera costing less than $500 (CV S3200 Rev B, JAI, Japan) since one purpose of this study was to determine if one could use an inexpensive camera with a limited sensitivity and dynamic range to monitor optical changes in tissue.

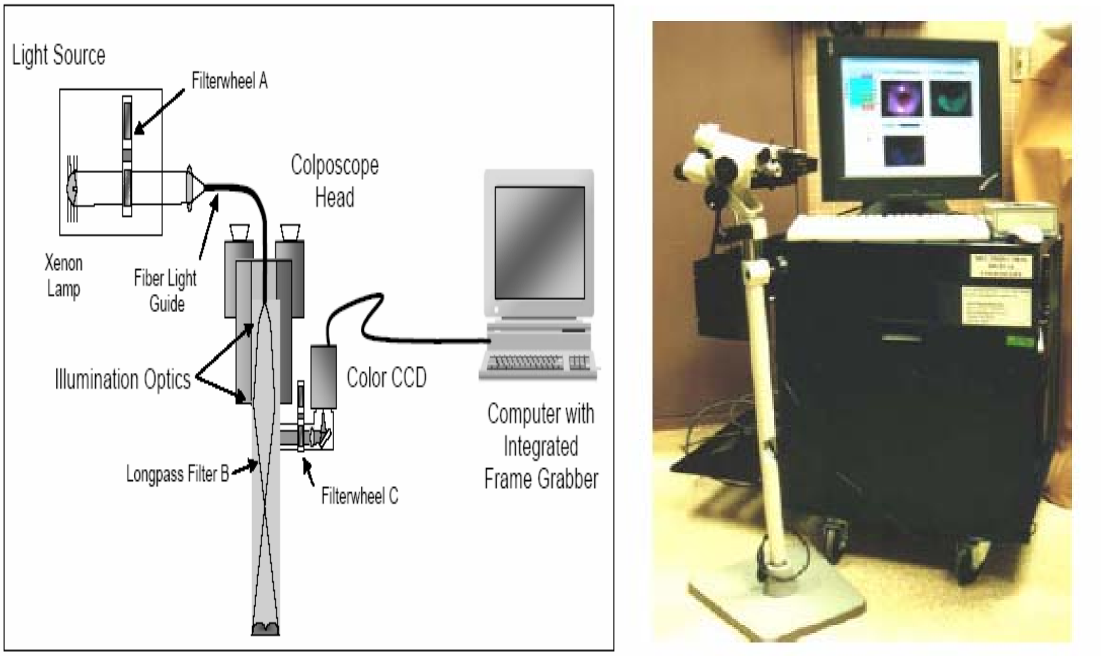


Figure 2. 1 (a) System diagram and (b) a photograph of the pilot MDC system [39]

The camera is an analog video rate color Cyan, Magenta, Yellow (CMY) camera. The dynamic range of each channel is 8 bits. The camera CCD is sensitive to CMY colors; however, the camera transforms the measured CMY values into Red, Green, Blue (RGB) values. The operation of the CCD camera was controlled using a PC computer and a LabVIEW program (National Instruments, Austin, TX).

2.4.2 New MDC

Based on the pilot MDC, a new MDC was designed. There are two major differences between the pilot and the new devices. The new MDC can provide cross and parallel polarized imaging in addition to unpolarized white light reflectance and fluorescence images. Polarized imaging has been augmented in order to reduce a specular reflection captured by the former pilot MDC. With unpolarized white light imaging, we found it difficult to obtain diagnostic information from areas obscured by specular reflection and this can potentially lead to errors in diagnosis. Polarized imaging can remove specular reflection as previously studied by Li [40]. Figure 2.2 shows an illustration for specular reflection removal through cross polarized imaging. An image measured using unpolarized white light (Figure 2.2(a)) shows areas of specular reflection, while an image measured for the same site using cross polarization (Figure 2.2(b)) does not contain specular reflection.

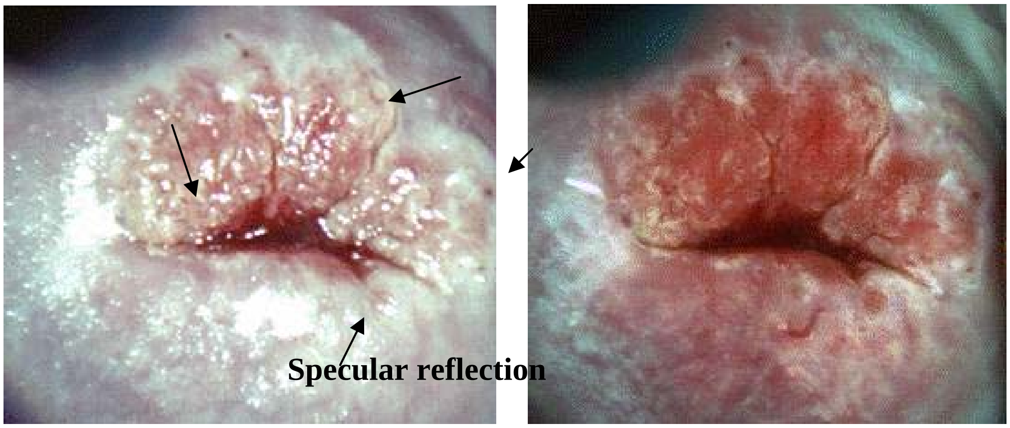


Figure 2. 2 (a) Unpolarized reflectance image (b) Cross polarized image

Furthermore, the new MDC is equipped with a higher resolution, higher dynamic range CCD camera which is more suitable for multi-spectral imaging than the CCD camera used for the pilot MDC. With the high resolution CCD camera, the new MDC can measure more diagnostic information than the pilot MDC, and, potentially, yield better diagnostic performance.

A picture of the new MDC system is shown in Figure 2.3. The new MDC was built based on a commercially available colposcope from Zeiss medical solutions (Model: KSK 150FC). The MDC light source is a powerful 120W Metal Halide lamp (X-Cite® 120) which provides monochromatic and broadband illumination. The excitation light generated from the light source is delivered to the tissue through a liquid light guide (Lumatec).

Figure 2. 3 A photograph of new MDC

The reflectance and fluorescence from tissue are captured and filtered using a commercially available digital color charge-coupled device (CCD) camera (QImaging MicorPublisher 5.0, USA) which provides 10-bit digital color images. Flexible exposure control from 1.6 ms to 17.9 min, high frame rates for previewing or focusing, minimal thermal noise during low-light imaging and high resolution (3.3 megamixel sensor) images are some of the advantages of this camera. Reflectance images are acquired under three different settings such as unpolarized white light, cross and parallel polarization. Fluorescence images are acquired at two excitation wavelengths UV (370nm with 10 bandwidth) and blue (440nm with 30 bandwidth) on the detection side, 430 nm for UV and 485 nm for blue. Fluorescence and polarized light are controlled using high speed automated filter wheels in the illumination path and the detection path.

A Pentium IV processor computer interacts with the excitation and detection sub-systems through a filter wheel controller, and provides a Graphical User Interface for both image acquisition and the quality control program.

2.5 DIAGNOSTIC IMAGE ANALYSIS

With an imaging system, diagnostic image analysis techniques are essential parts of field spectroscopy applications. Recently, various studies on image analysis techniques have been performed in order to assess the entire field of the cervix for the detection of neoplastic lesions [26-40]. In particular, automated image diagnosis algorithms based on modern image processing techniques have the potential to replace clinical expertise, which may considerably reduce the cost of developing a screening program in developing countries.

In the early 1990s, a number of studies investigated whether digital image processing techniques could be used to automate the interpretation of colposcopic images [26-38]. Craine et al. [27] introduced a digital colposcope system for archiving images and visually assessing features in the images. Shafi et al. [28] and Cristoforoni et al. [29] used a digital imaging system for colposcopy which enables image capture and simple processing. For assessment of various colposcopic features, the acquired images were manually analyzed by an expert. By examining the relationship between colposcopic features and histology outcomes, Shafi’s study provided information about which features are most useful to the expert observer. Image interpretation in these early studies mainly relied on experts’ qualitative assessment of colposcopic images and provided limited quantitative analysis.

Since these early studies, further automated algorithms have been designed with the goal of circumventing provider intervention. One of major diagnostic features used in these algorithms is aceto-whitening. Acetic acid enhances the differences in appearance between normal and dysplastic tissue. Figure 2.4 shows pre- and post acetic acid images measured through cross polarized imaging using the MDC system for a patient with CIN 2. The cervix of this patient has CIN2 and CIN3 at 5 and 6 o'clock position, and acetowhitening area was encircled by a black line in Figure 2.4(b).

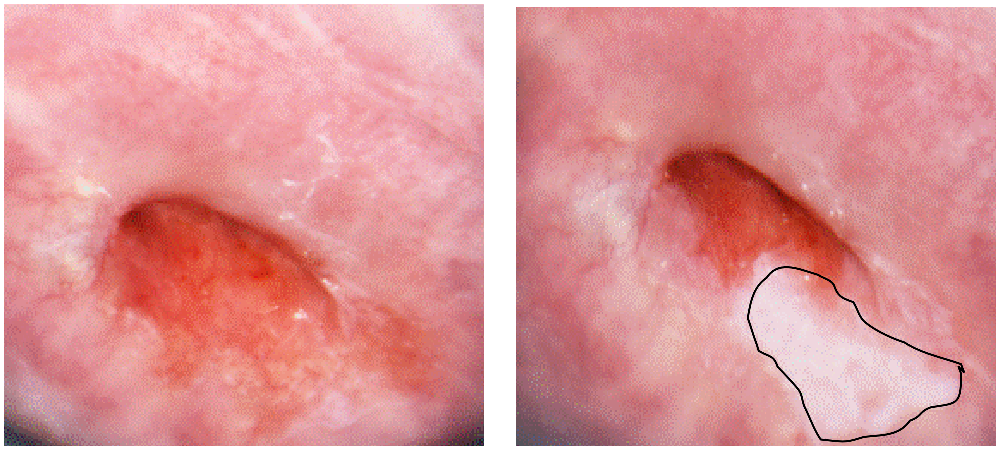


Figure 2. 4 (a) Pre acetic acid and (b) Post acetic acid cross polarized image

Li et al. [31] developed a computer-aided diagnostic system using colposcopic features such as acetowhitening changes, lesion margin and blood vessel structures. They prototyped image processing algorithms for detection of those features and showed promising preliminary results, however the diagnostic performance of the system has not yet been reported. Balas [33] developed a multispectral imaging system which is capable of performing time-resolved imaging spectroscopy. The temporal alterations in the intensity of backscattered light after acetic acid application enabled differentiation between neoplastic and non neoplastic lesions. The study showed diagnostic potential, but the diagnostic performance of this method was not reported. Gordon [34, 35] applied an unsupervised Expectation-Maximization (EM) clustering algorithm to separate acetowhite lesion from normal squamous epithelium and columnar epithelium using the color intensity of pixels and initial segmentation results were shown.

Another recent approach for colposcopic image analysis incorporates information about tissue type into a diagnostic framework. Cervical tissue can be mapped into regions characterized by a predominance of squamous tissue (ectocervix), columnar tissue (endocervix), and a transformation zone combining these tissue types as shown in Figure 2.5. Diagnostic algorithms can then be applied within these regions.

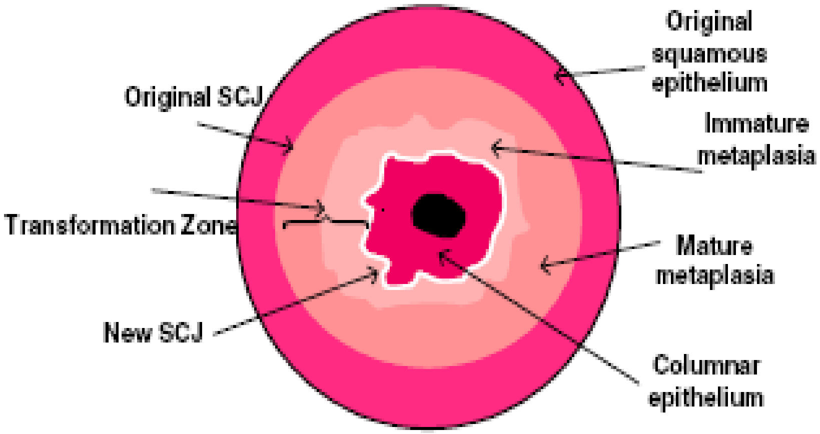


Figure 2. 5 Schematic presentation of a cervix [31]

Considering that cervical cancer is mainly caused by infection of metaplastic epithelium in the cervical transformation zone with one or more of carcinogenic types of human papillomavirus (HPV) [41], taking information about tissue type into account may increase the performance of the algorithm. Furthermore, the optical properties of tissue in the cervix vary by tissue type [19], and, therefore, delineating cervical tissue types in an image creates an opportunity to improve the performance of a diagnostic system by accounting for the imaging effects of varying cell structures.

Gordon [34] and Li [31] introduced an image analysis algorithm to segment the anatomical features of cervix such as columnar epithelium, squamous epithelium, the cervical canal and the transformation zone based on color intensity values. This research showed the potential of automatic segmentation of the transformation zone. However, any diagnostic feature analysis considering tissue areas was not reported and spatial relationship between tissue types was not considered. They also did not report the diagnostic performance of their approach.

In addition to image analysis techniques, there have been other challenges in studying image analysis related to how the diagnostic performance of image analysis algorithms can be reported. Many studies have reported their clinical performance based on biopsies or colposcopic impression. However, this approach cannot provide an accurate assessment for the diagnostic performance of the algorithms since neither biopsies nor colposcopic impression can yield a spatially varying diagnosis for the entire field of view of cervix. In order to report sensitivity and specificity for image analysis comparable to those for point spectroscopy analysis, a gold standard for the entire field of view is essential. Research presented in the dissertation directly addresses these limitations.

Chapter 3: Multispectral Digital Microscopy for In Vivo Monitoring of Oral Neoplasia in the Hamster Cheek Pouch Model of Carcinogenesis

This chapter has been published in the December 2005 edition of *Optics Express* [70].

3.1 INTRODUCTION

Animal models are of great importance in attempting to better understand the natural history and molecular biology of carcinogenesis. The hamster cheek pouch carcinogenesis model is a well-known and well-characterized animal model of human oral cancer and squamous cell carcinomas (SCC) [42, 43]. The model uses chronic treatments of the cheek pouch with DMBA. Histologically, the 16-week treatment protocol pushes the epithelial lining of the cheek pouch through stages of inflammation, hyperplasia, dysplasia, and both benign and malignant tumor formation [44, 45]. Epithelial hyperplasia develops after only a few treatments with DMBA. Dysplastic lesions, resembling human premalignant lesions, are seen after 6-8 weeks of treatment. After about 10 weeks, papillomas and carcinomas begin to appear [42].

Fluorescence spectroscopy is a promising new technology for in vivo detection of epithelial dysplasia [44-53]. A number of clinical trials of fluorescence spectroscopy have been performed in humans, demonstrating diagnostic potential in a variety of organ sites. To fully understand the diagnostic potential of fluorescence spectroscopy for early lesion detection, it would be useful to interrogate a single site throughout the entire dysplasia-carcinoma sequence. Obviously, ethical considerations prevent using fluorescence spectroscopy to follow the natural history of carcinogenesis in vivo in patients. As an alternative, several in vivo animal studies have been performed using fluorescence spectroscopy to follow the development of neoplasia in the hamster cheek pouch model [44, 45, 54]. Wang [55] measured autofluorescence emission spectra in vivo from DMBA-induced pre-cancers and early cancers in the hamster cheek pouch at 330 nm excitation. In the study, Wang showed results indicating that the autofluorescence spectroscopy technique is a useful diagnostic tool for in vivo diagnosis of oral pre-cancers and cancers in the DMBA-induced hamster cheek pouch carcinogenesis model. Vengadesan [56] studied native fluorescence spectra using the DMBA-induced hamster cheek pouch carcinogenesis model at 405 nm excitation, and characterized the native fluorescence of endogenous porphyrin and other fluorophores. Coghlan [44] showed that neoplastic lesions in the DMBA-induced hamster cheek pouch carcinogenesis model exhibit characteristic red fluorescence between 630-640 nm emission at 410 nm excitation, attributed to porphyrins. Coghlan also demonstrated that fluorescence spectroscopy could identify squamous intra-epithelial neoplasias with high sensitivity and specificity at optimal fluorescence excitation wavelengths of 350-370 nm and 400–450 nm [45]. Similar results were obtained in fluorescence measurements of early cancers of the human oral cavity in several clinical trials [46, 57-59]. Heintzelman [58] determined optimal excitation-emission wavelength combinations for in vivo identification of oral cavity neoplasia and showed that the optimal excitation wavelengths for detection of human oral neoplasia were 350, 380 and 400 nm. Muller [59] analyzed fluorescence data from oral tissue for the diagnosis of early oral carcinoma using trimodal spectroscopy. The results demonstrated the method is a highly sensitive and specific technique for tissue abnormality diagnosis.

Previous studies in the hamster cheek pouch model system have used a fiber optic probe to obtain emission spectra from a small area of tissue, typically with a radius less than 2 mm. Due to this limitation, it is difficult to screen the whole field at risk and to determine whether fluorescence can be used to effectively delineate the borders of neoplastic lesions. Recently, an MDM, which is a two dimensional, multispectral imaging system, has been developed and applied for the in vivo detection of cervical neoplasia [39]. Unlike the fiber optic spectroscopy approaches, this device has the advantage of being able to assess the entire field at risk.

In this study, we use the MDM to acquire multispectral fluorescence and reflectance images of the entire field at risk. The hamster cheek pouch model is adopted for the study since it provides a highly predictable model of neoplasia which can be studied over a short period of time. Images acquired at different time points in the study are analyzed to visualize and classify neoplastic lesions. In order to separate the chromatic information from the intensity information, the original data acquired in the RGB color space is transformed to the YCbCr color space. Various statistical parameters are extracted from the images such as the mean pixel value in each image channel, standard deviation, skewness, and kurtosis. A linear discriminant classification algorithm based on the Mahalanobis distance measure [60] is developed and tested using a cross validation technique. Results show that the chromatic information is an important parameter to identify the presence of neoplasia with high sensitivity and specificity.

3.2 MATERIALS AND METHODS

3.2.1 Multispectral digital microscope (MDM)

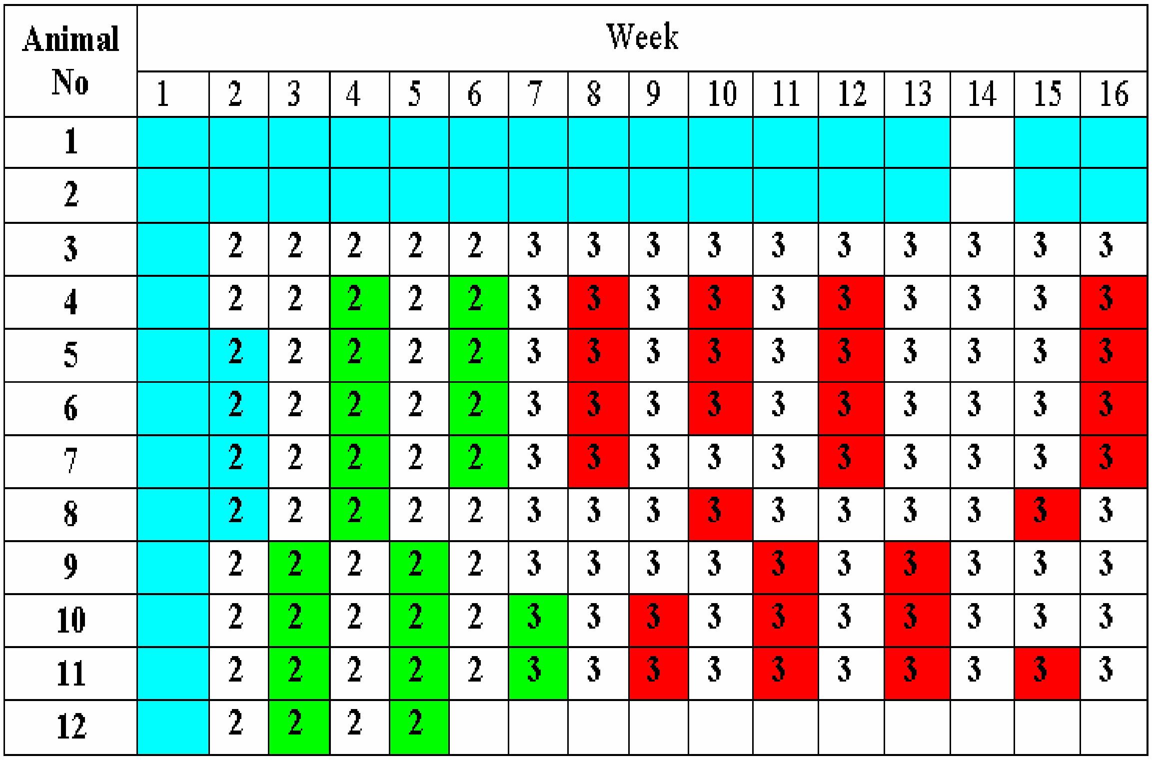
In order to assess the entire field at risk, an MDM was developed to measure multispectral reflectance images with white light illumination and autofluorescence images at 345 nm and 440 nm excitation. We chose these two excitation wavelengths based on an analysis of the two most useful wavelengths for diagnosis of cervical intraepithelial neoplasia from our previous hamster study and a large in vivo diagnostic trial of point spectroscopy [44, 45, 61].

The MDM is composed of a light source, a standard colposcope, and a video rate color CCD camera with a frame grabber [39]. A system diagram of the MDM system is shown in Figure 2.2(a). Details of this instrument can be found in chapter 2.

3.2.2 Study design

Table 1 provides an overview of the study design. The protocol was reviewed and approved by the M.D. Anderson Cancer Center Animal Care and Use Committee. A laboratory animal veterinarian with extensive experience with the DMBA model of hamster cheek pouch carcinogenesis (LC) participated in all aspects of the study. Twelve animals were followed weekly for 16 weeks. Ten animals were treated with the carcinogen 0.5% DMBA in mineral oil to induce gradual epithelial carcinogenesis, and two control animals were treated only with mineral oil. Coghlan [45] showed that treating with DMBA more than two times per week results in acute inflammation and significant erosion in the early weeks of the protocol. Therefore, animals were initially treated with carcinogen two times per week starting in the second week of the study in order to allow the inflammation to subside; after six weeks, treatments were increased to three times per week since the acute inflammation in the early weeks ceased. Given the potency of DMBA, there is little difference in the overall average incidence of lesions at 16 weeks.

Table 3. 1 Overview of the study design. DMBA weekly treatment frequency is indicated by the number in each cell; cells without a number indicate no DMBA treatment. Cells are color coded to indicate image acquisition and tissue classification. Different colors represent different tissue classifications: cyan for normal, green for intermediate, and red for neoplastic.



Animal No	Week															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦		🟦	🟦
2	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦		🟦	🟦
3	🟦	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3
4	🟦	2	2	🟩 2	2	🟩 2	3	🟥 3	3	🟥 3	3	🟥 3	3	3	3	🟥 3
5	🟦	🟦 2	2	🟩 2	2	🟩 2	3	🟥 3	3	🟥 3	3	🟥 3	3	3	3	🟥 3
6	🟦	🟦 2	2	🟩 2	2	🟩 2	3	🟥 3	3	🟥 3	3	🟥 3	3	3	3	🟥 3
7	🟦	🟦 2	2	🟩 2	2	🟩 2	3	🟥 3	3	3	3	🟥 3	3	3	3	🟥 3
8	🟦	🟦 2	2	🟩 2	2	2	3	3	3	🟥 3	3	3	3	3	🟥 3	3
9	🟦	2	🟩 2	2	🟩 2	2	3	3	3	3	🟥 3	3	🟥 3	3	3	3
10	🟦	2	🟩 2	2	🟩 2	2	🟩 3	3	🟥 3	3	🟥 3	3	🟥 3	3	3	3
11	🟦	2	🟩 2	2	🟩 2	2	🟩 3	3	🟥 3	3	🟥 3	3	🟥 3	3	🟥 3	3
12	🟦	2	🟩 2	2	🟩 2											

In each treatment, the treatment substance was applied to the deepest recess of the right cheek pouch with a No. 5 camel-hair brush. Images of the treated cheek pouch were acquired weekly with the MDM, except for week 14. Prior to the measurement, animals were anesthetized with 75-150 mg/kg BW ketamine HCI (Fort Dodge, Overland Park, KS 66210) delivered intraperitoneally, and the pouch was rinsed with saline solution. In order to expose the deepest recess of the pouch, the hamster cheek pouch was manually everted with a gloved index finger of the handler. At image acquisition, the hamster was placed in a lateral recumbent position within the handler's gloved hand.

The frame rate of the MDM is 30 frames/sec. Each white light image was taken at the full frame rate, while each flourescence image was obtained by summing 32 fields, resulting in a frame rate of 1.875 frames/sec. The working distance of the MDM was 30 cm. The measured illumination intensity at this working distance was approximately 0.25 mW/cm^2 at 345 nm excitation and 2.4 mW/cm^2 at 440 nm excitation. Fluorescence was acquired sequentially at the two excitation wavelengths. Approximately one minute was required to obtain a sequence of reflectance and fluorescence images.

Based on the well-established natural history of the DMBA hamster cheek pouch model, the measured images were divided into three groups: normal, intermediate, and neoplastic. The normal group included images from the two control hamsters and those of DMBA treated hamsters measured in week 1 before application of carcinogen and in week 2. Images of DMBA treated hamsters after week 7 were assigned to the neoplastic group. Images of DMBA treated hamsters from week 3 to week 7 were classified into the intermediate group, since their disease progress could not be determined precisely without a biopsy.

3.2.3 Preprocessing and color transformation

White light and fluorescence images acquired from the hamster cheek pouches were preprocessed for three purposes: noise reduction, image registration, and correction for variations in the illumination intensity. Each animal was imaged at two different fluorescence excitation wavelengths; slight motion between image acquisition at the different excitation wavelengths necessitated proper image registration. The intensity of illumination varied from day to day, leading to variations in overall image intensity which were unrelated to the disease process. Image preprocessing included the following steps. First, median filtering (3×3) was applied to each image to eliminate the fixed pattern noise due to the defective pixels in the CCD image sensor. Second, the two fluorescence images were aligned with the white light image. In this step, an automatic image registration algorithm based on mutual information was implemented to perform the image registration task without user intervention. Mutual information is a measure of the correlation of two images which is computed from the joint probability distribution of the two images' intensity. When two images are aligned, the resulting joint probability distribution of two images provides high-valued mutual information [62, 63]. Based on this property, white light and fluorescence images of the same object were aligned by maximizing the mutual information for each pair of these images. Finally, changes in the color and intensity of the image were separated using a color space transformation. Since the RGB color space is sensitive to both chromatic and intensity variations, images were transformed to the YC_bC_r color space, which achieves better color constancy [64]. The mathematical representation of this transformation [65] is provided in Eq. (1).

$$\begin{aligned} Y &= c_1R + c_2G + c_3B, \\ C_b &= \frac{B-Y}{2-2c_3}, \\ C_r &= \frac{R-Y}{2-2c_1}, \end{aligned} \tag{1}$$

where $c_1 = 0.2989$, $c_2 = 0.5866$, and $c_3 = 0.1145$.

3.2.4 Feature selection and classification

Once images were transformed to the YCbCr color space, simple statistical features were examined to determine whether they provided useful diagnostic information. A previous study [45] showed increased fluorescence near 350-370 nm excitation and decreased fluorescence near 450-470 nm excitation associated with neoplasia. Based on this observation, the mean image intensity was selected as a potential feature for classification. As neoplasia progressed, we observed that the intensity variation across an image increased, therefore the standard deviation was explored as another potential feature for classification. The intensity distribution of the normal data set was much more similar to a Gaussian distribution than corresponding distribution of the neoplastic data set. A common statistic for assessing the normality of a distribution is the Jaque-Bera (JB) test statistic, which is determined by skewness and kurtosis of the data [66]. Based on this fact, skewness and kurtosis were also identified as potential features for image classification.

A classification algorithm was developed to discriminate images as normal or neoplastic. Input features include the mean image intensity in each channel at each excitation wavelength, the corresponding standard deviation of pixel intensity in each image channel at each excitation wavelength, skewness and kurtosis, for a total of 24 input parameters. The classifier is a linear discriminant algorithm based on the Mahalanobis distance measure.

The classifier was trained using images from the normal and neoplastic groups. For training purposes, an expert small animal veterinarian (LC) was asked to circle the neoplastic regions in the corresponding white light image for each animal in the neoplastic group. For each of normal group images, 2-3% of the pixels in the 345 nm and 440 nm excited fluorescence images were randomly selected; for each of the neoplastic group images, 2-3% of the pixels in the encircled neoplastic regions of the 345 nm and 440 nm excited fluorescence images were randomly selected. The random sampling was performed using the uniform distribution. The statistical features associated with these randomly chosen pixels were used to train the classifier to identify tissue as normal or neoplastic. This random sampling technique was applied in order to reduce the likelihood of overtraining and to increase the robustness of the algorithm. The algorithm was trained using data from all but one hamster in the normal and neoplastic groups. This algorithm was then applied to data from the held-out animal and to animals from the intermediate group. In applying the algorithm, 2-3% of the pixels from the 345 nm and 440 nm excited fluorescence images were randomly chosen as input to the classifier without regard for the location of the areas encircled by the expert veterinarian. This process was repeated until all hamster images from the normal, intermediate, and neoplastic groups had been classified. Note that the classification results for the intermediate group provided additional information on diagnostic features even though they were not utilized when the quantitative diagnostic performance measures such as sensitivity and specificity were evaluated.

3.3 RESULTS

Figure 3.1 shows a time course of white light and autofluorescence RGB images of the DMBA treated cheek pouch of hamster 10. Fluorescence intensity changes associated with the dysplasia to carcinoma sequence can be clearly seen in Figure 3.1(b)-(e). The images in Figure 3.1 are representative of those obtained in the study. The left column shows the white light images. In each white light image, abnormal areas identified by the expert veterinarian are enclosed with a green line. The middle and right columns show fluorescence images acquired at 345 nm and 440 nm excitation, respectively. Images obtained in week 1, before application of carcinogen, serve as base line images for comparison with those in weeks 5, 7, 11, and 13. In week 5, a dysplastic lesion was observed. By week 7, a small exophytic neoplastic lesion has formed. These lesions progress to an obvious neoplasm in week 11. Consistent with human neoplasms, the malignant tumor in week 13 has begun outgrowing its blood supply and appears necrotic. At 345 nm excitation, dysplasia and tumor appear brighter than the

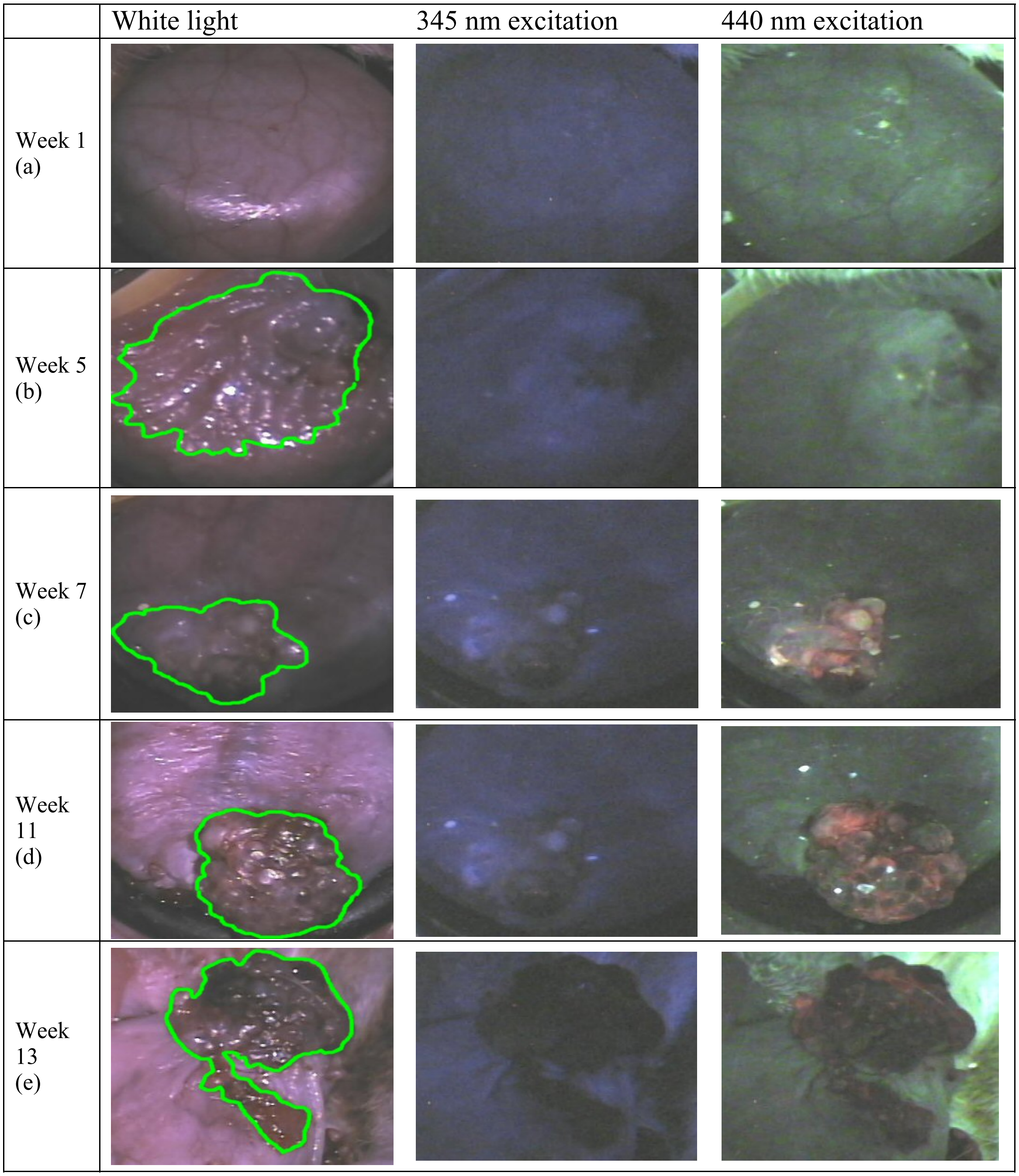


Figure 3. 1 Time course of RGB images for a representative DMBA treated hamster: a) 1st week, b) 5th week, c) 7th week d) 11th week, and e) 13th week (left: white light reflectance images, middle: 345 nm excited fluorescence images and right: 440 nm excited autofluorescence images).

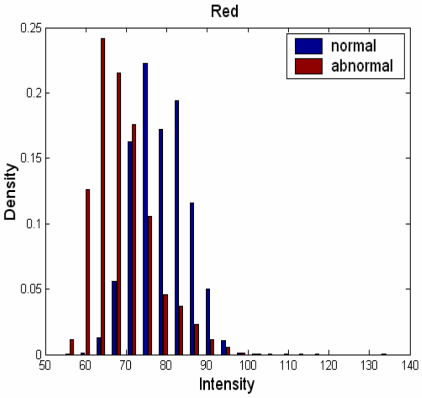
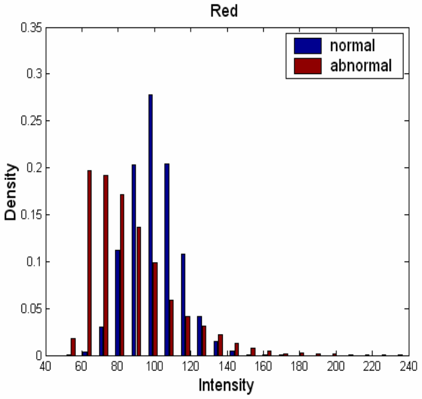
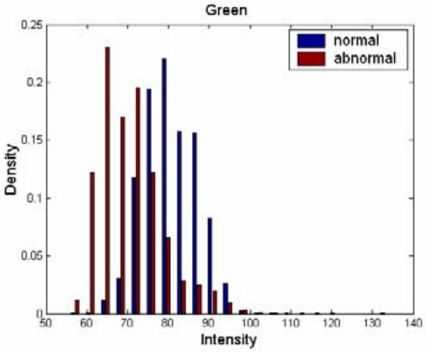
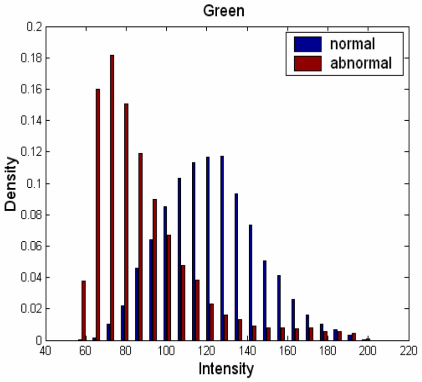
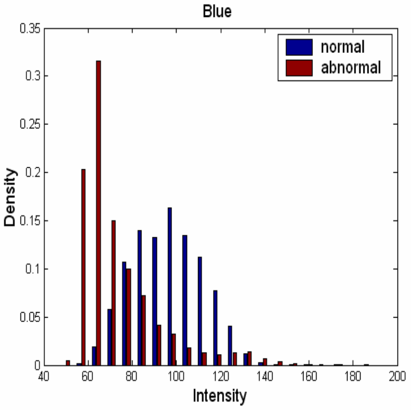
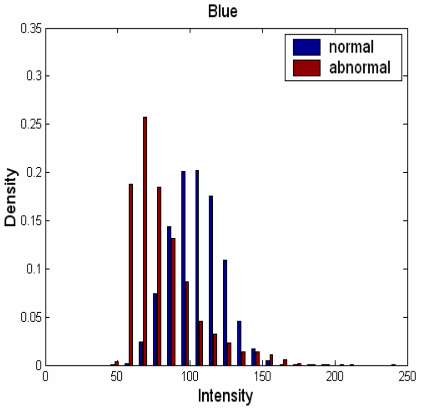
	345 nm (a)	**440 nm (b)**
R		
G		
B		

Figure 3. 2 Relative frequency histograms of the mean intensity of pixels randomly chosen from fluorescence images at (a) 345 nm and (b) 440 nm excitation in the normal and neoplastic groups. The left column shows data from the red channel, the middle column from the green channel and the right column from the blue channel.

surrounding normal tissue. However, the necrotic tumor at week 13 appears dark. At 440 nm excitation, the dysplasia looks lighter green. As the tumor appears and progresses, red fluorescence is observed in the images (Figure 3.1(c)-(e)). However, as shown in Figure 3.1(e), we consistently observed a decrease in the red fluorescence decreased when the malignant tumor formed in all DMBA treated hamsters.

Figure 3.2 shows relative frequency histograms of the parameters extracted from pixels randomly chosen from images in the normal and neoplastic groups. Figure 3(a) shows histograms of the mean intensity in the three color channels from the pixels randomly chosen from the normal and neoplastic group images in the RGB color space. Figure 3(a) shows histograms of the mean fluorescence intensity at 345 nm excitation for the red (left), green (middle), and blue (right) color channels. Figure 3(b) shows similar plots for the pixels from images acquired at 440 nm excitation. At both 345 nm and 440 nm excitation, the histogram of the mean intensity of the pixels of images from the abnormal hamsters is asymmetric and shifted to lower values, while that from the normal hamsters is fairly symmetric and shifted to higher values.

Figure 3.3 (a), (b), (c), and (d) show scatter plots of the mean, the standard deviation, the skewness, and the kurtosis, respectively, of the intensity at 345 nm excitation versus values at 440 nm excitation for each hamster at each time point in the normal, intermediate and neoplastic groups. Results are shown for each of the red, green, and blue channels. A noticeable separation between the normal and the neoplastic data can be observed in the mean intensity plots for the red, green, and blue channels. Interestingly, the mean intensities from images in the intermediate group are clustered in between those of the normal and neoplastic groups. The intermediate group represents a transition from normal to dysplasia, suggesting that the mean intensity feature changes continuously as neoplasia progresses. However, the other statistical

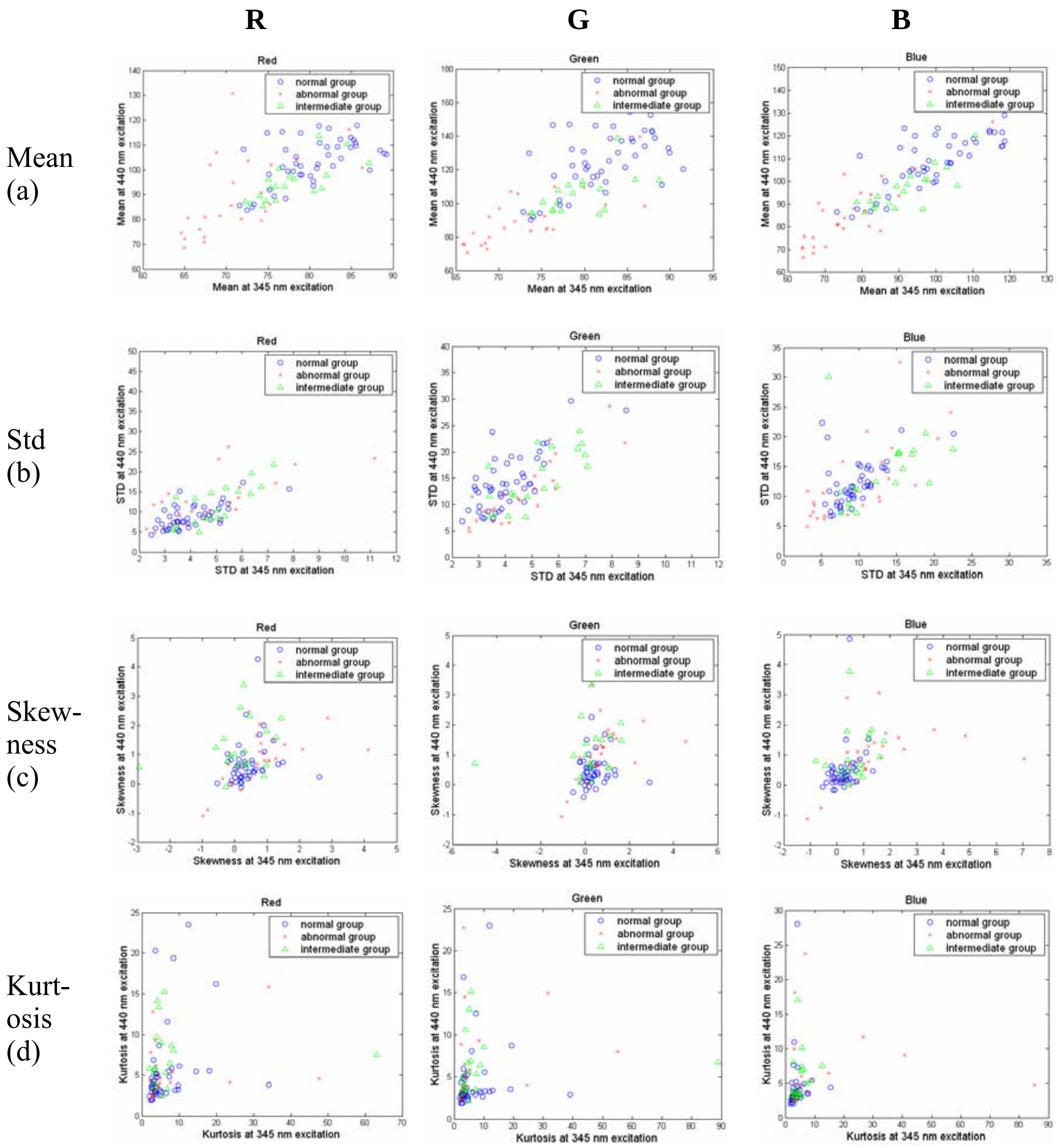


Figure 3. 3 Two dimensional scatter plots of the statistical parameters from pixels randomly chosen from fluorescence images at 345 nm excitation vs the same parameters at 440 nm excitation, including (a) mean intensity, (b) standard deviation, (c) skewness and (d) kurtosis. The left column shows data from the red channel, the middle column from the green channel and the right column from the blue channel. A data point is shown for each hamster at each time point in the normal, intermediate and neoplastic groups.

measures do not show any clear separation in this color space. Using these RGB parameters as input, a linear discriminant classifier based on the Mahalanobis distance measure was applied; images from the normal and neoplastic groups were classified with a sensitivity of 84% and a specificity of 79%.

Figure 3.4 shows similar relative frequency histograms for the mean intensity, standard deviation, skewness and kurtosis in the YC_bC_r color space of pixels randomly chosen from images in the normal and neoplastic groups. Figure 3.4(a) shows the mean intensity histograms at 345 nm excitation for the Y (left), C_b (middle), and C_r (right) channels. Figure 3.4(b) shows the mean intensity histograms plots at 440 nm excitation. At 345 nm excitation, the mean intensity histograms of the neoplastic group are asymmetric and shifted to lower values in both the Y and C_b channels, while that in C_r channel is shifted to higher values. At 440 nm excitation, the mean intensity distribution in the Y channel is asymmetric and shifted to lower values in the neoplastic group, while slightly asymmetric and shifted to higher values in both the C_b and C_r channels. At both excitation wavelength and for all three channels, the mean intensity distributions of the normal hamsters are symmetric.

Figure 3.5 (a), (b), (c) and (d) show two dimensional scatter plots of the mean, the standard deviation, the skewness, and the kurtosis of the intensity at 345 nm excitation versus those at 440 nm excitation, respectively, for each hamster at each time point in the normal, intermediate and neoplastic groups for each of the Y, C_b, and C_r channels. Clear separation between the normal and the neoplastic data is observed in both the C_b and C_r channels for all statistical parameters except kurtosis. In the Y channel, no clear separation is observed for any of the statistical measures. Using these YC_bC_r parameters as input, a linear discriminant classifier based on the Mahalanobis distance measure was applied; images from the normal and neoplastic groups were classified with a sensitivity of 96% and a specificity of 84%.

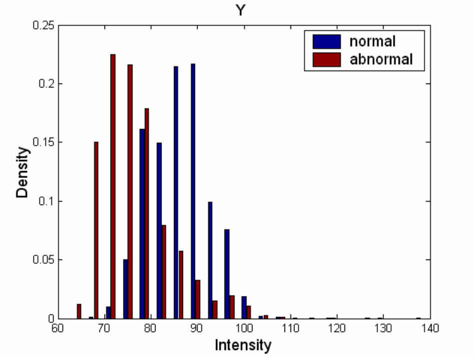
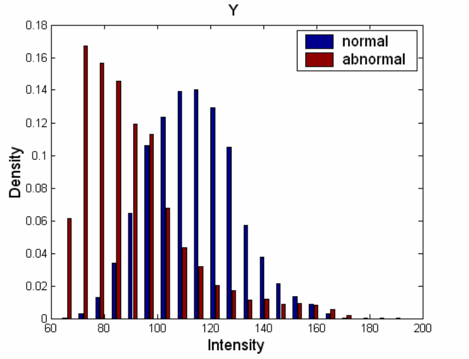
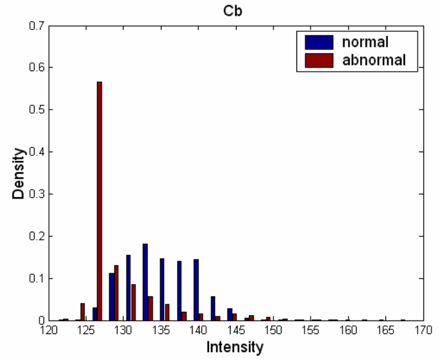
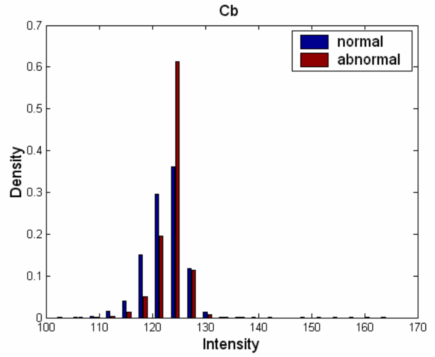
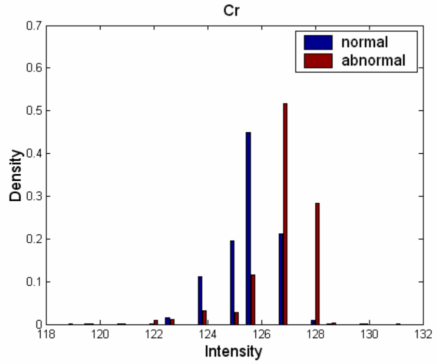
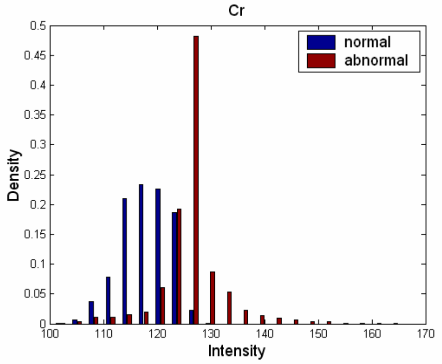
	345 nm (a)	**440 nm (b)**
Y		
C_b		
C_r		

Figure 3. 4 Relative frequency histograms of the mean intensity of pixels randomly chosen from fluorescence images at (a) 345 nm and (b) 440 nm excitation in the normal and neoplastic groups. The left column shows data from the Y channel, the middle column from the C_b channel and the right column from the C_r channel.

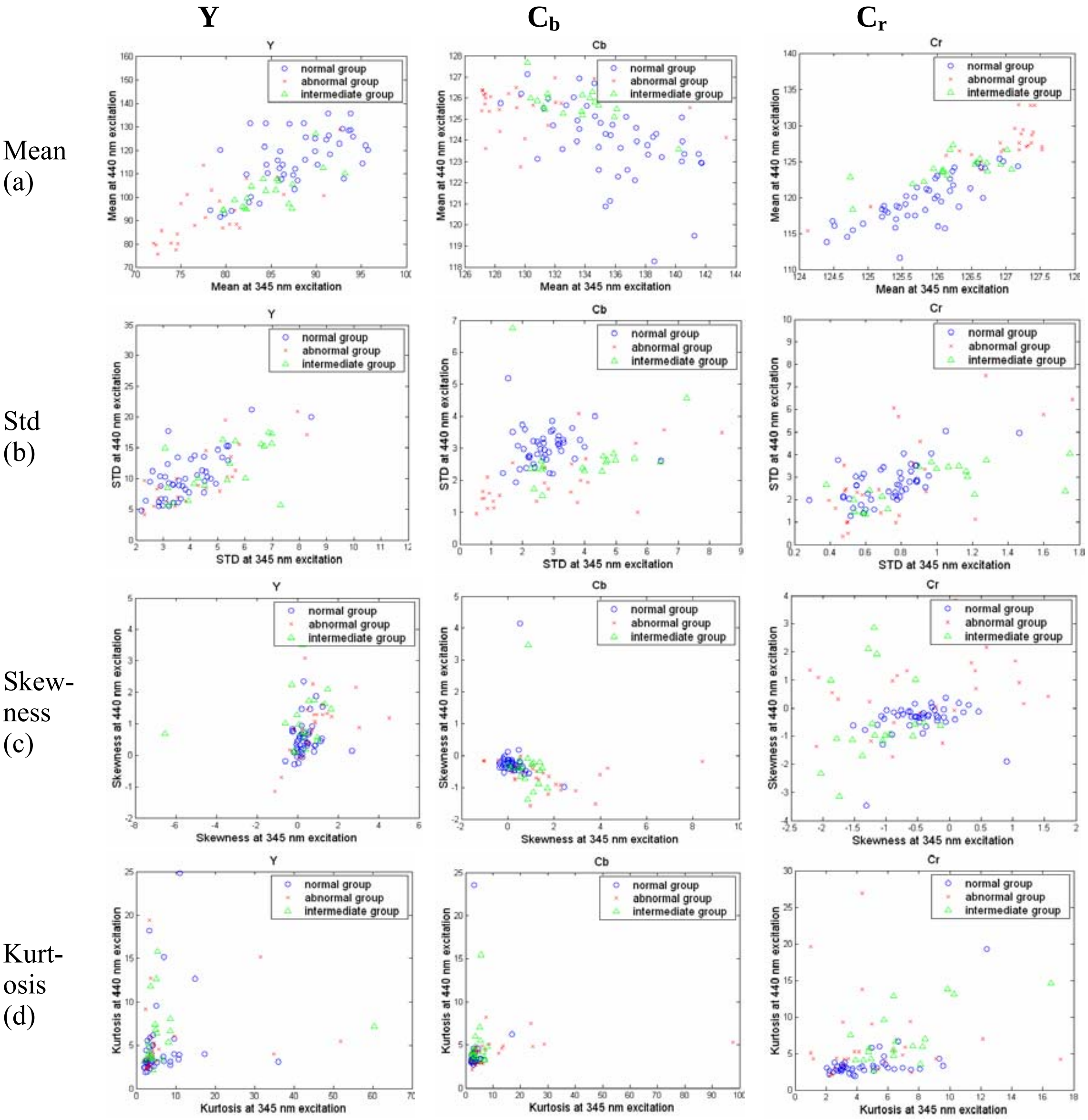
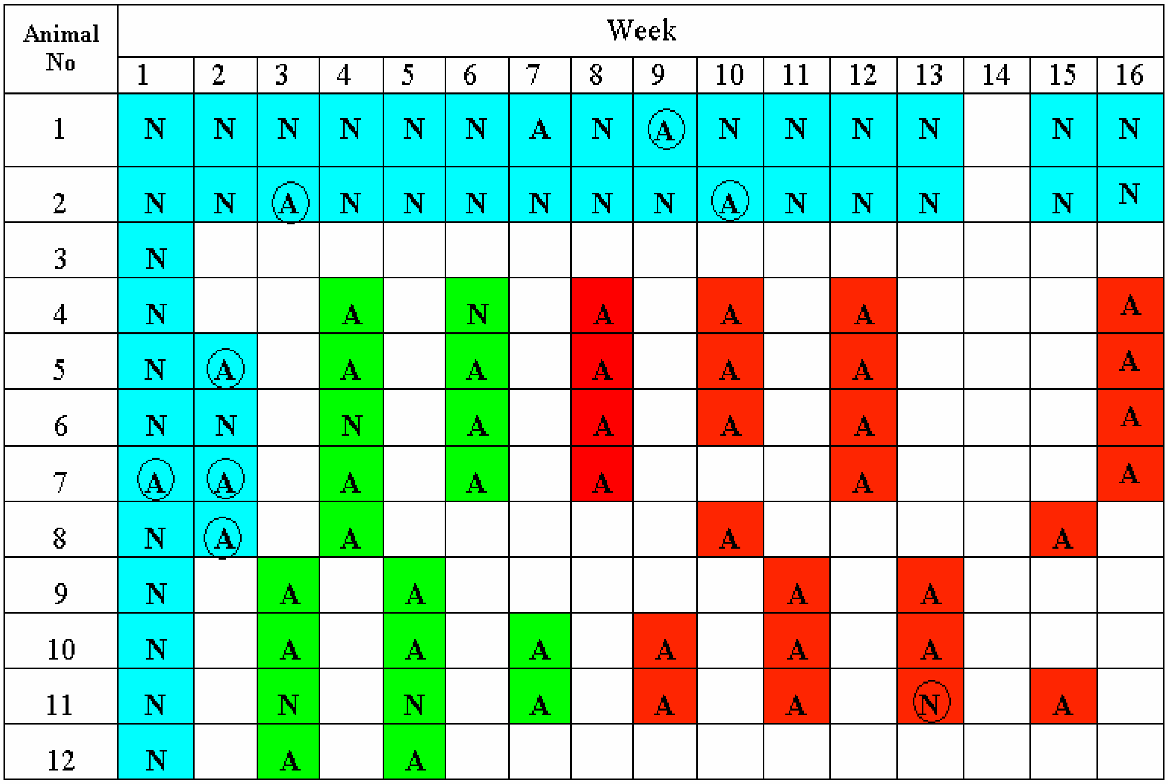


Figure 3. 5 Two dimensional scatter plots of the statistical parameters from pixels randomly chosen from fluorescence images at 345 nm excitation vs the same parameters at 440 nm excitation, including (a) mean intensity, (b) standard deviation, (c) skewness and (d) kurtosis. The left column shows data from the Y channel, the middle column from the C_b channel and the right column from the C_r channel. A data point is shown for each hamster at each time point in the normal, intermediate and neoplastic groups.

Table 3. 2 Classification results using the input data from images in the YC_bC_r color space for each hamster at each time point. Colored cells indicate time points at which images were acquired. The cells are color-coded to indicate which group the images were assigned to as in Table 3.1. The results of the classification algorithm are represented by the letter in the cell, where N indicates the measurement was classified as normal and A indicates the measurement was classified as neoplastic. Circled letters indicate misclassifications for hamsters in the normal and neoplastic groups.



Animal No	Week															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	N	N	N	N	N	N	A	N	Ⓐ	N	N	N	N		N	N
2	N	N	Ⓐ	N	N	N	N	N	N	Ⓐ	N	N	N		N	N
3	N															
4	N			A		N		A		A		A				A
5	N	Ⓐ		A		A		A		A		A				A
6	N	N		N		A		A		A		A				A
7	Ⓐ	Ⓐ		A		A		A				A				A
8	N	Ⓐ		A						A					A	
9	N		A		A						A		A			
10	N		A		A		A		A		A		A			
11	N		N		N		A		A		A		Ⓝ		A	
12	N		A		A											

Table 3.2 presents the detailed results of the classification algorithm applied to images in the YC_bC_r color space for each hamster at each time point. Only one image from the neoplastic group was misclassified. Although the images from the intermediate group

were not included in the calculation of sensitivity and specificity, it is interesting to examine the classification results for this group. Most of the images in this category were classified as abnormal. Furthermore, for all but one hamster, once data from an animal was classified as abnormal, the images at subsequent time points were also classified as abnormal.

3.4 DISCUSSION

Compared to point spectroscopy, multispectral imaging with the MDM has the advantage of providing a real time assessment of the entire field at risk. In this study, we show that a simple analysis extracting image features such as the mean image intensity or standard deviation can yield diagnostic performance which is similar to that of point spectroscopy [45]. Best results are obtained when the images are represented in the YC_bC_r color space, enabling good separation of intensity and chromatic changes. In particular, moving to the YC_bC_r color space simplifies the ability to monitor the increase in red fluorescence associated with neoplasia at 440 nm excitation. This observation has been reported in previous spectroscopy studies, and is attributed to porphyrin fluorescence. The origin of this porphyrin fluorescence is controversial and whether it is associated with tumor cells or with microbial contamination is unclear [67]. In the RGB color space, due to the high correlation of all three RGB color channels with intensity variations [64], the changes in red fluorescence are mingled with changes in overall intensity. In the YC_bC_r color space, the intensity variations are clearly separated from changes in chromatic features. Earlier studies with point spectroscopy indicate that the presence of red fluorescence is associated with neoplasia, but is not always present. In this study, an increase in red fluorescence was consistently observed at 440 nm excitation near the neoplastic lesions for all the DMBA treated hamsters. However, the red fluorescence was heterogeneously distributed, and studies with point spectroscopy that interrogate only a small point on the tissue may not adequately sample the lesion. The MDM provides a more holistic assessment of the field at risk. In the future, it may be possible to further improve diagnostic performance by exploring spatial variations in statistical parameters throughout the image.

A potential weakness of this study is that a confirmatory biopsy was not obtained at each time point. While it would provide definitive diagnostic information, such frequent biopsies would clearly suppress the development of neoplasia and would significantly alter the natural history of the lesions. Since the hamster cheek pouch model is so well-characterized with clear time points at which neoplasia develops, the use of the model with the lack of biopsies was adopted rather than performing frequent biopsies. In summary, the MDM provides an inexpensive tool to monitor changes in the autofluorescence associated with neoplasia. Use of such an inexpensive system may prove useful for improved cancer screening in regions with limited health care resources. Clinical applications of this technique may include improved early detection, better direction of diagnostic biopsy and real time determination of tumor margins.

Chapter 4: White Light Reflectance Image Analysis Using Acetowhitening Feature

4.1 INTRODUCTION

Cervical cancer is the second most common cancer in women worldwide, and 80% of cases occur in the developing world where the fewest resources exist for management [1]. Most cases of cervical cancer can be prevented through screening programs aimed at detecting precancerous lesions. Screening for cervical neoplasia using the Papanicolaou (Pap) smear, followed by colposcopy, biopsy, and treatment of neoplastic lesions has dramatically reduced the incidence and mortality of cervical cancer in every country in which organized programs have been established [68]. This is due to the fact that the precursors, denoted as Cervical Intraepithelial Neoplasia (CIN) or Squamous Intraepithelial Lesions (SIL) can take three to twenty years to develop into cancer. However, due to the lack of resources and infrastructure, 238,000 women die each year of cervical cancer; more than 80% of these deaths occur in developing countries [1, 4]. We are interested in applying optical technologies to replace expensive infrastructure for cervical cancer screening in the developing world.

The use of Direct Visual Inspection (DVI), Visual Inspection with Acetic Acid (VIA), and Visual Inspection with Lugol's Iodine are being explored as alternatives to Pap smear and colposcopic examination in many developing countries [7-10]. Recent reviews of the performance of these methods report that they have sufficient sensitivity and specificity, when performed by trained professionals, to serve as viable alternatives to Pap screening in low-resource settings [11-13]. Table 4.1 shows the results of a review by Sankaranarayanan, et al., which favorably compares these methods to HPV testing and conventional cytology [14].

Table 4. 1 Results of Pooled Analyses from Sankaranarayanan [14].

Screening Methodology	**No. of Studies Reviewed**	**No. of patients**	**Percent of patients with HG and greater**	**Sensitivity pooled**	**Sensitivity Range**	**Specificity Pooled**	**Specificity Range**
VIA	11	54,981	16	0.77	0.56-0.94	0.86	0.74-0.94
VIAM	3	16,900	14	0.64	0.61-0.71	0.87	0.83-0.90
VILI	10	49,080	16	0.92	0.76-0.97	0.85	0.73-0.91
Cytology	5	22,663	6	0.58	0.29-0.77	0.95	0.89-0.99
HPV HC II	4	18,085	7	0.67	0.46-0.81	0.94	0.92-0.95

VIA is visual inspection with acetic acid, VIAM is visual inspection with acetic acid and magnification, VILI is visual inspection with Lugol's iodine, Cytology means conventional cytology, and HPV HC II is human papillomavirus hybrid capture II test (Digene, Gaithersburg, Maryland)

Because DVI relies on visual interpretation, it is crucial to define objective criteria for the positive identification of a lesion and to train personnel to correctly implement a program based on these criteria. Denny noted that restricting the definition of a positive VIA test to a well-defined acetowhite lesion significantly improved specificity, while reducing sensitivity [13]. In a series of 1,921 women screened in Peru, Jeronimo found that the DVI positivity rate dropped from 13.5% in the first months to 4% during subsequent months of a two year study; the drop in positivity rate was hypothesized to be due to a learning curve for the evaluator [15]. Bomfim-Hyppolito investigated the use of cervicography as an adjunct to DVI. A simple Sony digital camera was used to photograph the cervix before and after application of acetic acid [13]. Photographs were later interpreted by an expert colposcopist. The addition of cervicography improved both sensitivity and specificity. However, this approach prevents the implementation of see-and-treat strategies in low resource settings because of the need for expert review.

Recently, optical techniques have been investigated as an alternative detection method in a quantitative and objective manner. Several studies have demonstrated that optical spectroscopy has the potential to improve the screening and diagnosis of neoplasia. Ferris et al. studied multimodal hyperspectral imaging for the noninvasive diagnosis of cervical neoplasia [22]. They reported a sensitivity of 97% and a specificity of 70%. Huh et al. measured the performance of optical detection of HGSIL using fluorescence and reflectance spectroscopy, finding a sensitivity of 90% and a specificity of 70% [23].

As another promising application of optical techniques for cervical cancer screening, a number of studies investigated whether digital image processing techniques could be used to automate the interpretation of colposcopic images [26-29]. Craine et al. [27] introduced a digital colposcopy system for archiving images and visually assessing features in the images. Shafi [28] and Cristoforoni [29] used a digital imaging system for colposcopy which enables image capture and simple processing. For assessment of various colposcopic features, the acquired images were manually analyzed by an expert. By examining the relationship between colposcopic features and histology outcomes, Shafi's study provided information about which features are most useful to the expert observer. Image interpretation in these early studies mainly relied on experts' qualitative assessment of colposcopic images and provided limited quantitative analysis. Li et al. developed a computer-aided diagnostic system using colposcopic features such as acetowhitening changes, lesion margin and blood vessel structures [31]. They prototyped image processing algorithms for detection of those features and showed promising preliminary results. However, the diagnostic performance of the system has not been reported.

Recently, advances in consumer electronics have led to inexpensive, high dynamic range Charge-Coupled Device (CCD) cameras with excellent low light sensitivity. At the same time, advances in vision chip technology allow high quality image processing in real time. These advances may enable acquisition of diagnostically useful digital images of the cervix in a relatively inexpensive way, with or without magnification. Moreover, automated analysis algorithms based on modern image processing techniques have the potential to replace clinical expertise, which may reduce the cost of screening. The purpose of this study is to explore whether digital colposcopy, combined with recent advances in camera technology and automated image processing, could provide an inexpensive alternative to Pap screening and conventional colposcopy.

Several challenges for diagnostic digital colposcopic image analysis still remain. (1) Previous studies have investigated only a few features and have not taken advantage of the mature field of image analysis and computer automated techniques. (2) Previous studies have compared image features of selected normal and abnormal areas of the cervix, but have not applied the approach to an entire image to identify whether lesions are present. (3) Previous studies have used biopsies from selected areas as the gold standard. A gold standard is needed for the entire field of view in order to address the issue of lesion localization.

In this pilot study, we explored the potential of an approach to automated image analysis of white light colposcopy images that addresses the three challenges discussed above. First, our approach employed automated image analysis techniques including image registration, pattern recognition, clustering, and classification. Second, our algorithm can identify pre-neoplastic tissue areas from an entire image. Third, a gold standard for the entire cervical image was constructed using a whole cervix specimen acquired from a Loop Electrosurgical Excision Procedure (LEEP) which was intensively sectioned.

4.2 MATERIALS AND METHODS

4.2.1 Instrumentation

A multispectral digital colposcope (MDC) was developed to acquire reflectance images of the entire cervix with white light illumination. The MDC consists of a commercially available, tilt-stand colposcope (Model 1DL, Leisegang, Germany) with a video rate color CCD camera and frame grabber for image acqusition. Details of this instrument can be found in papers by Benavides et al. [39] and Park et al. [70]. The colposcope produces stereoscopic vision at 7.5X magnification. Reflectance images were captured using an inexpensive, commercially available, video rate, color CCD camera (CV S3200 Rev B, JAI, Japan).

4.2.2 Clinical Study

The study protocol was reviewed and approved by the institutional review boards at The University of Texas M. D. Anderson Cancer Center, The University of Texas Health Science Center in Houston, The Lyndon B. Johnson County Hospital and the Harris County Health District, The British Columbia Cancer Agency, Rice University, and The University of Texas at Austin. Eligible patients were at least eighteen years old, not pregnant, and were referred to the colposcopy clinic with an abnormal Papanicolaou smear. Written informed consent was obtained from each participant, and all patients underwent a history, complete physical exam, Papanicolaou smear, cultures for gonorrhea and chlamydia, pan-colposcopy (including cervix, vagina, vulva, and perianal area), and colposcopically directed biopsies. Only patients with high-grade lesions who were scheduled at a previous visit were eligible for the study. Following colposcopic examination, but prior to the treatment with a LEEP, white-light reflectance images were acquired with the MDC from each patient at baseline. A second set of reflectance images was acquired following the application of acetic acid (6%) on the cervix for two minutes. Acetic acid enhances the differences in appearance between normal and dysplastic tissue. The patients then underwent a LEEP procedure. The histopathology from the specimen was inked and cut by the study pathologist. Each ectocervix was cut into twelve pieces. The authors reviewed all the hematoxylin and eosin stained slides jointly, noting areas of cervical intraepithelial neoplasia (CIN). Most specimens are extensively sectioned. The histopathologic slides were scanned and the images were uniformly enlarged by 15% to account for shrinkage by formalin [71], under the assumption of equal shrinkage. The images were then reconstructed into a three dimensional "pathology map" delineating areas of intraepithelial neoplasia using 3D image Visualization Software® (Able Software, Lexington, MA). Pathology maps were reviewed by the study clinicians and pathologists and compared to the white light reflectance MDC images.

4.2.3 Data Processing

Image pre-processing was performed to remove impulse noise spikes, specular reflection, and regions obscured by the presence of blood on the tissue surface. Median filtering was utilized for removing impulse noise spikes. For specular reflection removal, pixels with extreme intensity values above a certain threshold value (>220) were removed. Blood regions were identified by comparing red channel intensity with the average red channel intensity of the whole tissue regions, and removed. In order to properly register images obtained before and after acetic acid application, an automatic registration algorithm was developed. In the registration algorithm, the location of the cervical os was used as a reference point.

Our approach to automated digital image analysis for screening for cervical neoplasia can be divided into two major steps. First, tissue regions with similar optical properties are clustered together. Second, classification algorithms are used to determine whether these regions contain neoplastic tissue. These steps are described in detail below.

4.2.4 Classification of Localized Image Areas

To identify optical parameters which are most useful to determine whether localized image areas contain neoplastic tissue, the original red-green-blue (RGB) color intensity feature space was extended and features for image analysis were selected from this extended feature space. The extended feature space is composed of five categories of features: 1) raw RGB intensity values, 2) ratios of individual RGB intensity values, 3) differences in individual RGB intensity values, 4) differences in RGB intensity ratios, and 5) changes in grayscale intensity values following application of acetic acid. For each of the pre- and post-acetic acid images, the ratios of all possible pairs of features were calculated. Acetic acid-induced differences in RGB intensity values, ratios, and grayscale values were also calculated.

Features for classification were selected using the joint entropy maximization principle. The crucial assumption of this approach is that features with more information will result in better classification results. Since joint entropy represents the total information from a set of features, by maximizing joint entropy, a set of features valuable for classification can be selected. In this study, Shannon entropy was used and the joint entropy is defined as:

$$H(A_1, A_2, \cdots, A_n) = -\sum p(a_1, a_2, \cdots, a_n) \log_2 p(a_1, a_2, \cdots, a_n),$$

where $A_1, A_2, \cdots, A_n$ are random variables representing feature values of a pixel for an image application, $a_1, a_2, \cdots, a_n$ are realizations of $A_1, A_2, \cdots, A_n$, respectively, and $p(\bullet)$ is the joint probability mass (density) function of $A_1, A_2, \cdots, A_n$.

However, evaluating the joint probability mass function of more than three variables takes a large amount of computation time. In order to avoid this computational burden, the following assumption was made:

$$\text{If } H(A_1, A_2, \cdots, A_m) > H(A'_1, A'_2, \cdots, A'_m),$$
$$\text{then } \sum_{\substack{i,j \in \{1,\cdots m\} \\ i \neq j}} H(A_i, A_j) > \sum_{\substack{i,j \in \{1,\cdots m\} \\ i \neq j}} H(A'_i, A'_j),$$

where m is the number of selected features and $\{A_1, A_2, \cdots, A_m\}$ and $\{A'_1, A'_2, \cdots, A'_m\}$ are two different feature sets from the total feature set of an image $\{A_1, A_2, \cdots, A_n\}$. Under this assumption, the sum of joint entropies for all possible pairs of features has been calculated instead of the total joint entropy. Using the method described above, five features were selected as potentially useful for classification: intensity values for red, green, and blue channels, the ratio of intensity in the green channel to that in the red channel, and the changes in grayscale intensity values.

Given these five features, we next developed classifiers to estimate the tissue type of a localized region. We designed an ensemble classifier, wherein a number of different classifiers are applied and then the final classification is based on the most highly predictive of the individual classifiers. The four classifiers in the ensemble are a linear classifier with Euclidian distance, a linear classifier with Mahalanobis distance, a K-nearest neighbor (KNN) classifier with eight neighbors, and a support vector machine with linear kernel.

The classifiers were trained using leave-one-patient-out cross-validation [72]. In this method, the algorithm is trained on all data excluding one patient and then applied to the held-out patient; then the process is repeated successively for all patients in the data set. It is very useful in small studies, but may overtrain the algorithm. For each patient, image regions were identified corresponding to histopathologically proven areas of normal epithelium or low grade squamous intra-epithelial lesions (LGSILs) and high grade squamous intra-epithelial lesions (HGSILs). Because HGSIL is treated and LGSIL is followed, HGSILs were considered to be abnormal, while areas of normal tissue and LGSIL were considered to be normal. Training data were obtained using a window-based approach wherein 20×20 pixel windows were selected from areas of normal tissue or minor atypia, LGSIL, and HGSIL within an image. For each window, the five features were calculated for each pixel. Then the mean, the standard deviation, the 95th percentile, and entropy of each feature were calculated for the 20×20 pixel window.

The 20×20 pixel windows forming the training set were sampled in three different ways to assess the effect of training data on the resulting classifier performance. In the first approach (training data set 1) abnormal areas were selected from regions which were both pathologically HGSIL and in which the acetic acid intensity change was visually obvious. In the second approach (training data set 2), abnormal areas were selected from pathologically HGSIL regions in which the acetic acid intensity change was visually less apparent. In the third approach (training data set 3), abnormal data were selected based only the pathology map and without reference to the colposcopic images. The classifier performance was tested using the method of cross-validation and evaluated using receiver-operator characteristic curve (ROC) analysis. The classifier designed with training set data was then used to classify entire images as described below.

4.2.5 Image Analysis

The previous section described how features useful for classification were identified, and how classifiers were developed for local image regions. However, for truly automatic image interpretation, procedures must be developed to automatically segment the image into locally similar regions.

Local regions for analysis were segmented using a variant of the K-means clustering algorithm. In order to overcome the well-known local minima problem of the K-means algorithm, we employed a variant based on perturbation theory. The perturbation probability is determined for each mean based on its distance. We define the perturbation probability p_i for the cluster *i*'s mean by:

$$p_i = \frac{\underline{d}_i}{\sum_{k=1}^{K} \underline{d}_k},$$

where $\underline{d}_i$ is the distance between the mean of cluster *i* and its nearest mean. Note that the value of the perturbation probability changes with every iteration of the algorithm.

At every iteration of the algorithm, one of the K means is selected for perturbation in a probabilistic way using the perturbation probability. Suppose that the cluster *i*'s mean is selected to be perturbed by the perturbation probability at the νth iteration. Then, the perturbed mean of cluster *i*, $m'_{i\nu}$, is determined as follows:

$$m'_{i\nu} = m_{i\nu} + \Delta m_{i\nu},$$

where $\Delta m_{i\nu}$ is the random deviation for the perturbation which is selected from the normal distribution such that $\Delta m_{i\nu} \sim N\left(0, \sigma^2_{i\nu}\right)$. In this study, $\sigma_{i\nu}$ is designed as $\frac{\underline{d}_i}{1.96\nu}$ such that the variance of the perturbation decreases as the iteration proceeds, and, when the cluser *i*'s mean is selected at the first iteration, the magnitude of the deviation will be between 0 to $\underline{d}_i$ with the probability of 95%.

After segmenting the image into local regions using the cluster algorithm, image classification was performed using the algorithm developed in the previous section. For each cluster, windows of pixels were randomly selected from the cluster and, based on the classification results, a probability of abnormality was calculated and assigned to the cluster. The abnormal regions were then defined as those clusters for which the probability of abnormality exceeds a preset threshold value.

4.2.6 Performance Assessment

In order to assess the classification results of each image, a grid-based approach for calculating sensitivity and specificity was used. In this approach, an image is divided into grids and the classification results of the pixels in each grid of an image are compared with the corresponding pathology map. If the number of misclassified pixels is less than 20% of the number of pixels in the grid, then the grid point is considered to be correctly classified. Using this approach, the sensitivity, Se_{image}, and the specificity, Sp_{image}, of an image are defined as follows:

$$Se_{image} = \frac{number\ of\ grids\ correctly\ identified\ as\ positive\ grids}{number\ of\ positive\ grids},$$

$$Sp_{image} = \frac{number\ of\ grids\ correctly\ identified\ as\ negative\ grids}{number\ of\ negative\ grids}.$$

In order to assess the overall sensitivity and specificity of the automated image analysis procedure, the threshold probability resulting in the most desirable sensitivity and specificity values was determined in the Pareto sense [73]. This threshold probability was applied to all patients' image classification results. If an image contains an area of HGSIL whose radius is greater than 1mm, then the patient is classified as needing further treatment. Using this approach, the sensitivity, Se_{mdc}, and specificity, Sp_{mdc}, are defined as follows:

$$Se_{mdc} = \frac{number\ of\ patients\ correctly\ identified\ as\ HGSIL}{number\ of\ HGSIL\ patients},$$

$$Sp_{mdc} = \frac{number\ of\ patients\ correctly\ identified\ as\ LGSIL}{number\ of\ normal\ or\ LGSIL\ patients}.$$

4.3 RESULTS

Pre- and post-acetic acid white light images were measured from twenty-nine patients. Histopathology examination revealed twenty-one patients (73%) with CIN 2 or 3 lesions in the ectocervix, two patients (6%) with CIN 1 lesions, and six patients (21%) with no evidence of CIN. These eight patients each had a previous biopsy showing HGSIL.

Figure 4.1 illustrates the training data acquisition. As illustrated in the figure, abnormal data for the training set was sampled from pixels in HGSIL regions, and normal data for the training set were sampled from histopathologically normal or LGSIL regions.

Figure 4.2 illustrates the change in the ratio of the intensity in the green channel to that of the red channel induced by acetic acid for training set data identified as histologically normal and abnormal. The application of acetic acid does not substantially change the green to red intensity ratio measured from normal tissue, as shown in both the histogram and the box plot. However, the green to red intensity ratio increases for abnormal samples following application of acetic acid, as shown in both the histogram and the box plot. Following application of acetic acid, this parameter shows considerable separation between the normal and abnormal samples, and this intensity ratio was one of the features selected for use in the classification algorithm.

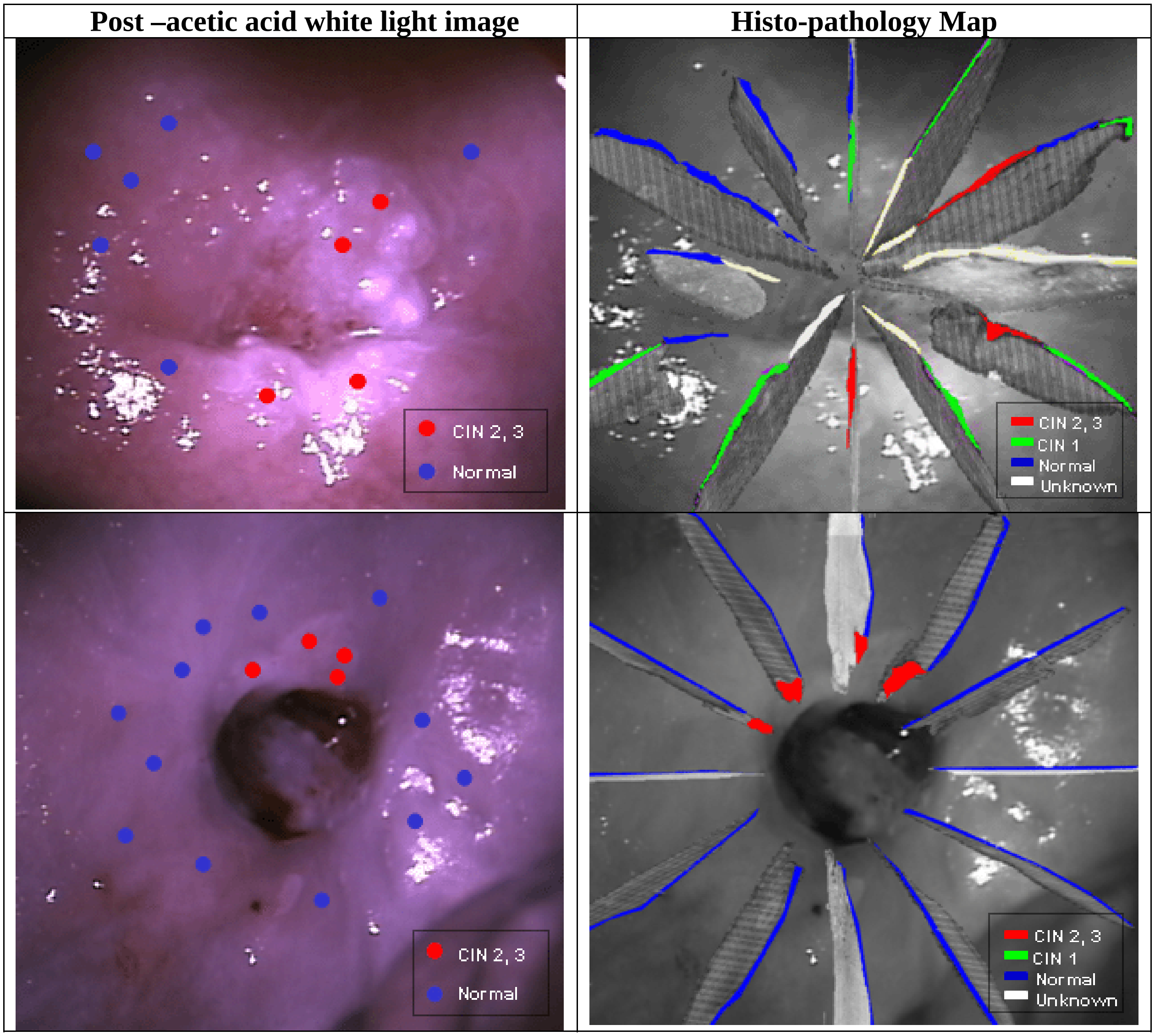
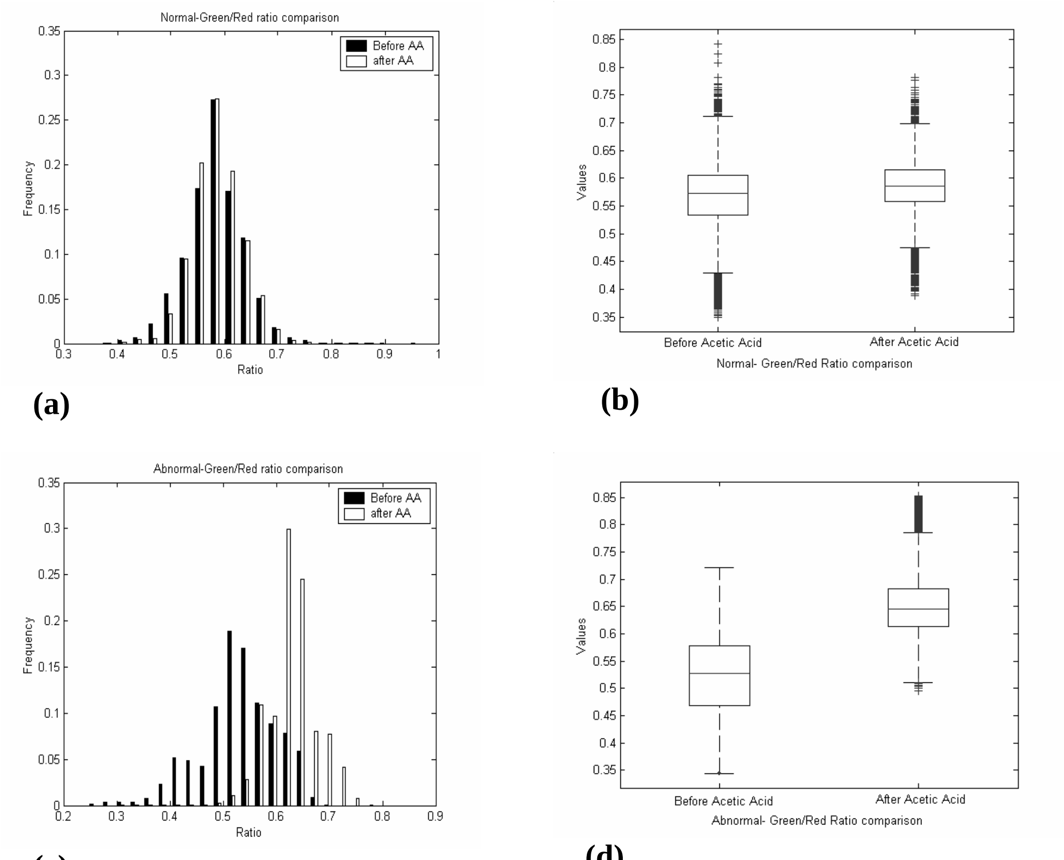


Figure 4. 1 Post-acetic acid white light images from two patients (left column) with uperimposed histopathology map (right column). The dots in the left column indicate the regions randomly selected for use in the localized image area classification. The tissue type at these points was determined according to the hisopathology map. The histopathology map includes an image of an H&E stained histopathology section of the LEEP specimen obtained at each clock position of the LEEP specimen. The epithelium within these sections is color coded according to the histopathologic diagnosis, with normal epithlieum indicated in blue, CIN 1 in green and CIN 2 and CIN 3 in red. Areas of missing epithlieum are colored in white.

Normal Samples

Abnormal Samples



(a) **(b)**

(c) **(d)**

Figure 4. 2 Histogram and box plot showing the ratio of the reflectance in the green channel to that in the red channel pre- and post-acetic acid. The (a) histogram and (b) box plot of the histologically normal samples indicates little change in this feature, while the (c) histogram and (d) box plot of this feature increase dramatically post-acetic acid for histologically abnormal specimens.

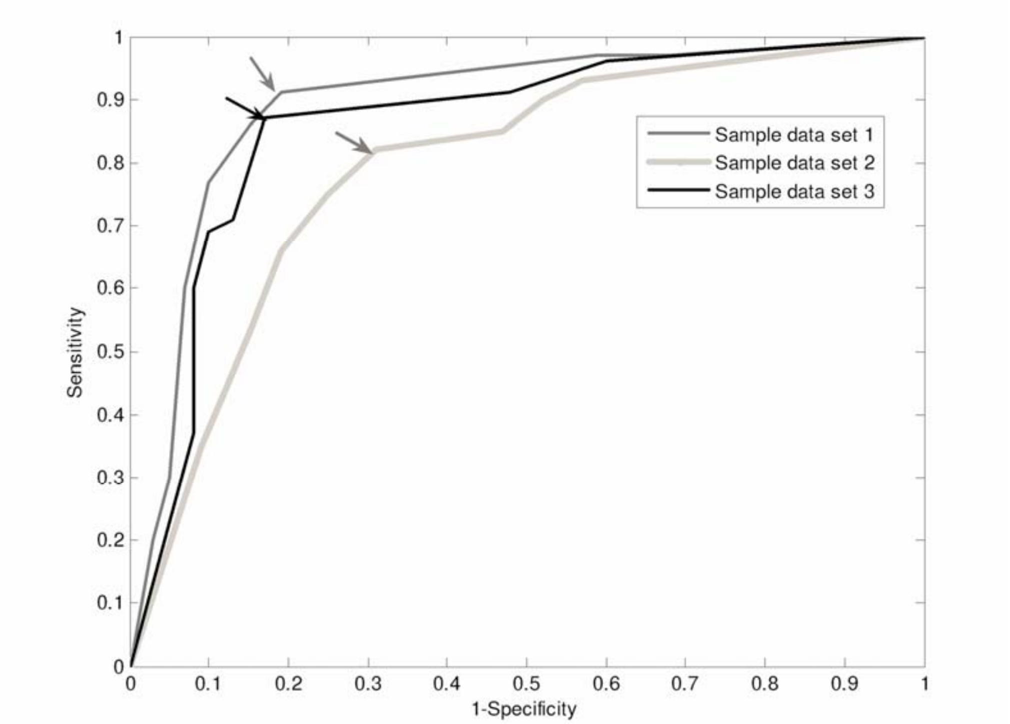
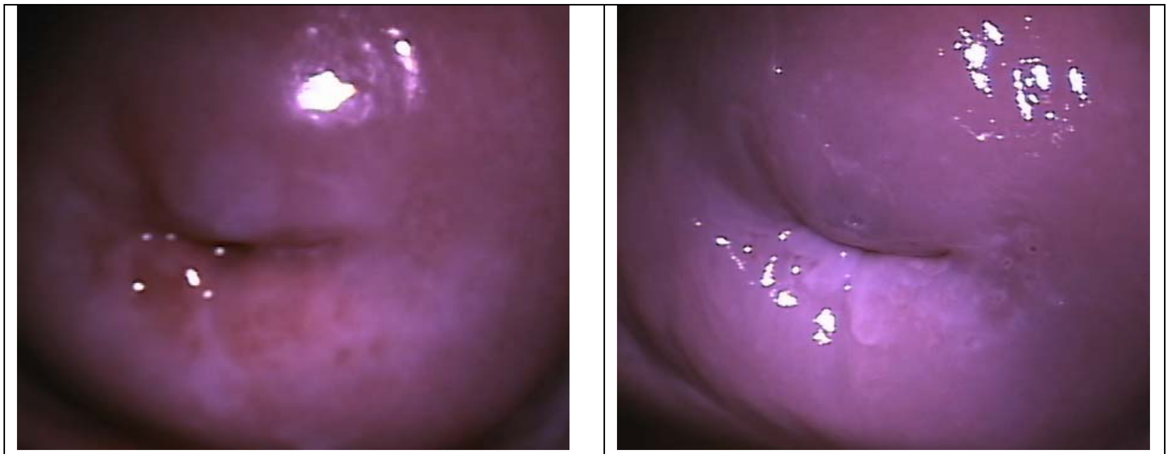
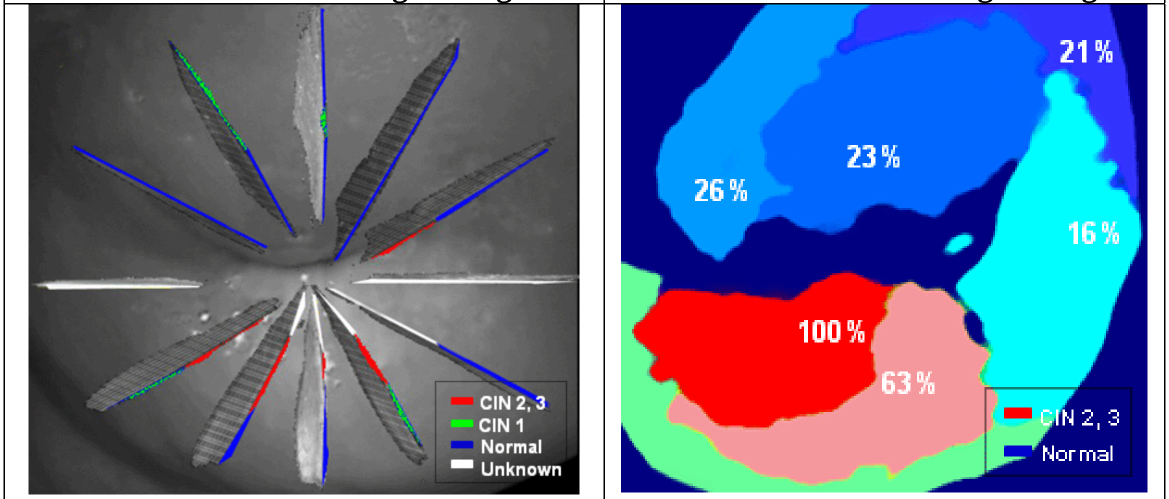


Figure 4. 3 ROC (Receiver Operating Characteristic) curves for multiclassifier discrimination of high grade SIL vs. normal or low grade SIL sample for three different sample data sets.

The performance of the proposed multi-classifier for localized image areas with three different training data sets has been analyzed and the resulting ROC curves are shown in Figure 4.3. As shown in Figure 4.3, training data set 1 yielded the best results; these data were sampled from those regions where both the pathology map indicated high grade disease and the colposcopic images showed significant changes after addition of acetic acid. The performance of training data set 3, where image regions were chosen based only on the pathology map, is better than that of training data set 2, where abnormal areas were selected from pathologically abnormal regions in which the acetic acid intensity change was visually less apparent. For each ROC curve, the point with the most desirable sensitivity and specificity values has been selected among Pareto-optimum points [73]. The selected point for training data set 1 corresponds to a sensitivity of 91% and a specificity of 80%. The optimum point for training data set 2 provided a sensitivity of 82% and a specificity of 69%, and that of training data set 3 corresponds to a sensitivity of 87% and a specificity of 83%.



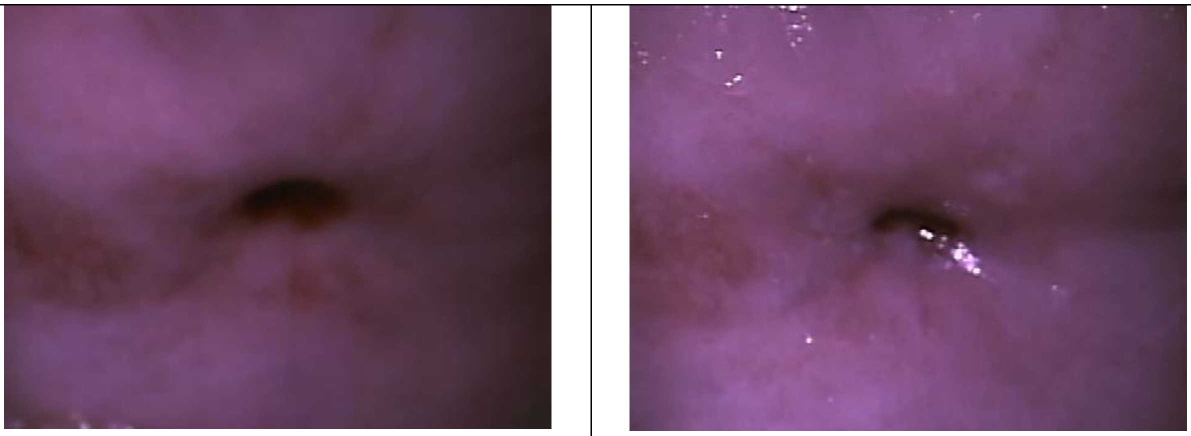
Pre-acetic acid white light image **Post-acetic acid white light image**



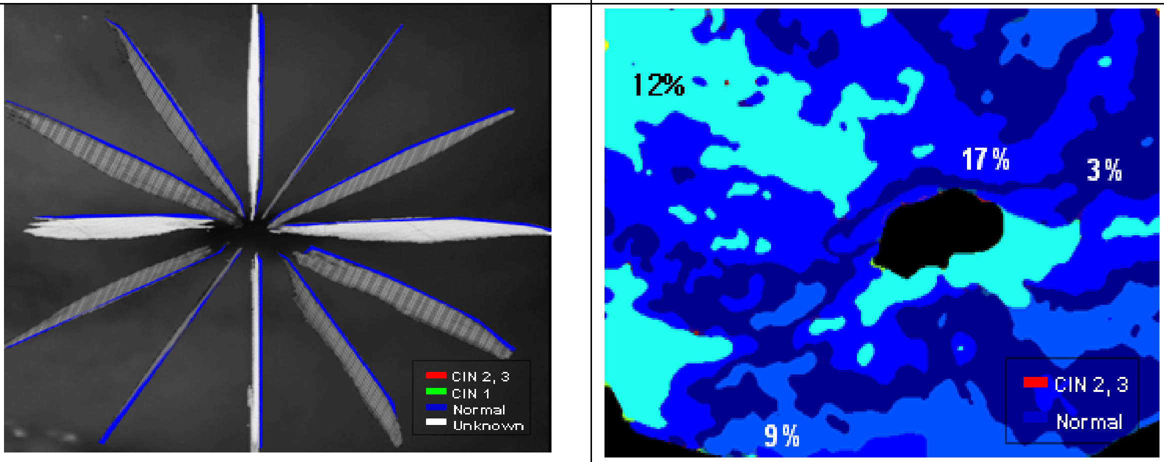
Histopathology map **Disease probability map**

Figure 4. 4 Image classification results for a patient with CIN 1, 2, and 3 lesions using the proposed image classification approach

Figure 4.4 shows pre- and post-acetic acid images for a patient with CIN 1, 2, and 3, together with the pathology map and the image classification result. The high-probability regions in the disease probability map correlate well with regions of CIN 2 and 3 in the histopathology map.



Pre-acetic acid white light image **Post-acetic acid white light image**



Histopathology map **Disease probability map**

Figure 4. 5 Image classification results for a patient with normal ectocervix using the proposed image classification approach

Figure 4.5 shows similar data for a patient with a completely normal ectocervix. The highest probability of abnormality in the probability map is 17% and, for any threshold value greater than 17%, all regions are diagnosed as normal. Figure 4.6 shows the performance of image classification for twenty-nine patients using ROC curve. The classification result of each patient was assessed using a grid-based approach. In the ROC curves, kink points are observed since there are only a limited number of probabilities of abnormality for clusters. That is, as we increase the threshold probability, the diagnostic performance changes in an abrupt manner only when the threshold probability passes the probability of abnormality for a cluster. Using the average ROC curve, the threshold probability (0.68) with the most desirable average sensitivity (82%) and specificity (73%) values was determined in the Pareto sense [74].

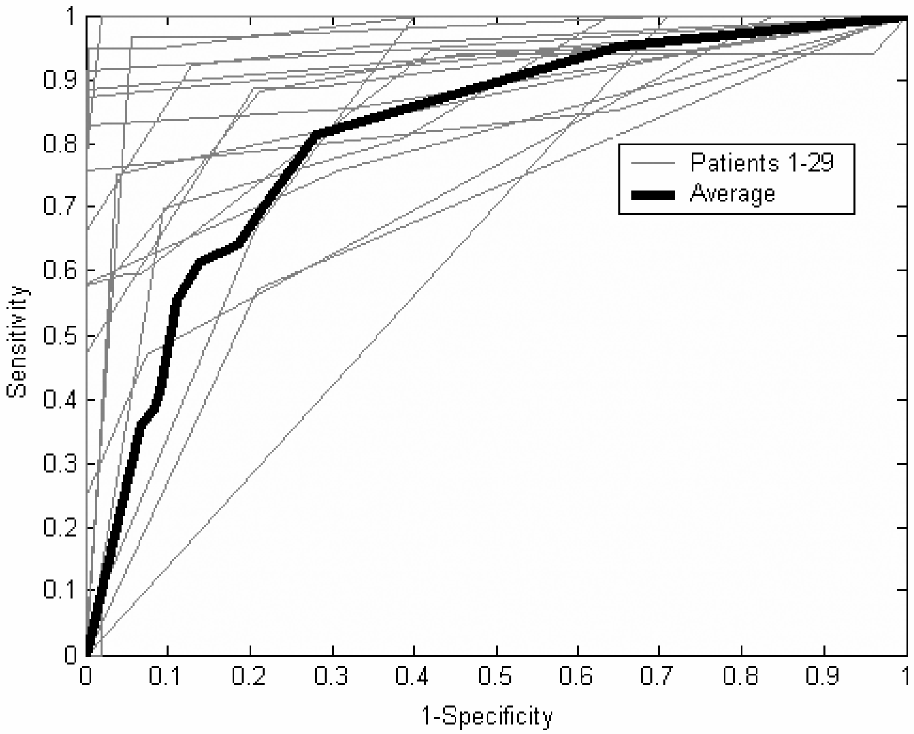
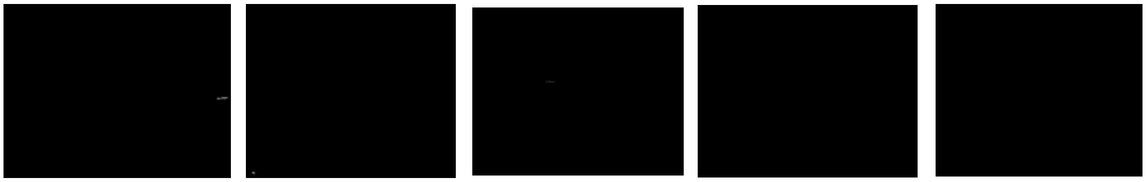
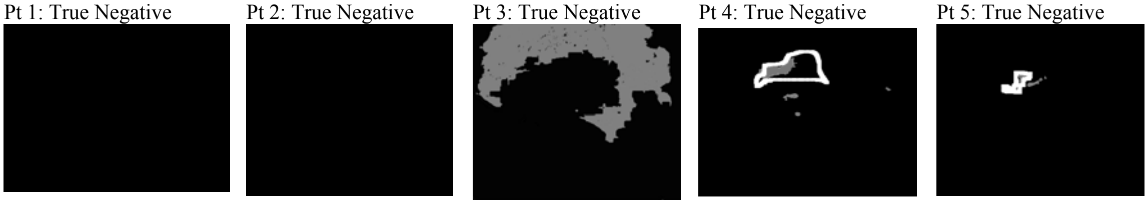


Figure 4. 6 ROC curves for 29 patients' image classification. The best threshold value (0.68) was assessed from the point of the average ROC curve corresponding to 82% of sensitivity and 73% of specificity.



Pt 1: True Negative Pt 2: True Negative Pt 3: True Negative Pt 4: True Negative Pt 5: True Negative



Pt 6: True Negative Pt 7: True Negative Pt 8: False Positive Pt 9: True Positive Pt 10: True Positive

Pt 11: True Positive Pt 12: True Positive Pt 13: True Positive Pt 14: True Positive Pt 15: True Positive

Pt 16: True Positive Pt 17: True Positive Pt 18: True Positive Pt 19: True Positive Pt 20: True Positive

Pt 21: True Positive Pt 22: True Positive Pt 23: True Positive Pt 24: False Negative Pt 25: False Negative

Pt 26: False Negative Pt 27: False Negative Pt 28: False Negative (Correct Decision) Pt 29: False Negative (Correct Decision)

Figure 4. 7 Automated image based diagnosis results for 29 patients with a normal epithelium or LGSIL or HGSIL within the ecto-cervix by LEEP. A threshold was applied to the disease probability map and those regions diagnosed as normal are represented in black, while the areas diagnosed as high grade SIL are shaded in gray. The area enclosed by a white line represents the area reported as CIN II or III from the pathology map. The algorithm yielded 7 true negative results, one false positive result, 15 true positive results and 6 false negative results.

Figure 4.7 shows the results of the automated image diagnosis after applying the threshold probability (0.68) for twenty-nine patients with a normal or LGSIL or HGSIL histopathology map. Patients #1-8 had a normal or LGSIL histopathology map (8 patients). For patients #1-7, the diagnosis based on automated image analysis matches the pathology report. However, in patient #8, there is a discrepancy between the automated image analysis result and the pathology report; the automated image analysis identifies a considerable abnormal area, but the pathology map indicates that there is no abnormal tissue present.

Patients #9-29 shown in Figure 4.7 demonstrated to have CIN 2 or 3 by LEEP (twenty-one patients). For fifteen patients where the classification algorithm yielded a true positive result, the area of HGSIL identified from the classification algorithm correlates well with that from the histopathology map (patients #9-23). Had the algorithm been used to determine whether treatment was indicated, it would have resulted in a correct decision. For six patients, the classification algorithm yielded an incorrect result (patients #24-27) or inconclusive result (patients #28 and #29). In four of these patients, the automated image analysis procedure did not identify areas of HGSIL later identified in the LEEP. In two patients, areas of HGSIL were identified in the image analysis procedure, but they did not correspond to the locations where HGSIL was later documented in the histopathology map. The resulting sensitivity and specificity values are 79% and 88%, respectively, for twenty-seven patients not including the two inconclusive results.

4.4 DISCUSSION

This pilot study shows promising agreement between our white light image analysis results and histopathology with a sensitivity of 79% and a specificity of 88%.

This is comparable to the performance of conventional colposcopy where the sensitivity and specificity have been reported to range from 87 – 96 % and 34 – 85 % [74-76].

Several other image processing approaches have been explored to aid in colposcopic image interpretation. Intensity change due to acetic acid application has been utilized in many diagnostic image analysis studies. Pogue, et al [30] analyzed acetowhitening of high-grade squamous intraepithelial lesions of six patients with biopsy confirmed CIN 2/3. They showed the temporal and spatial changes of green to red intensity ratio as an important acetowhitening feature to separate cervical lesions from normal regions using time course reflectance images. However, they did not report diagnostic performance using this feature. Balas et al. [33] assessed the alteration in the light scattering properties of the cervix using the maximum contrast between acetic acid responsive and non-responsive areas for sixteen patients. The time course of the intensity of backscattered light was adopted as an important diagnostic feature. The study showed the diagnostic potential of this technique; however, the quantitative diagnostic performance of the method was not reported. In our study, we developed a gold-standard based approach to rigorously assess sensitivity and specificity in the image classification context. This allows us to quantitatively assess the performance of our system.

A potential weakness of this study is that it was performed on patients with high-grade lesions; the predictive value may be lower in a screening population. In addition, our proposed approach may fail to detect precancerous lesions which do not show acetowhitening. There may be lesions that do not have keratin and thus are not aceto-white. We suspect that glandular lesions, both in situ and invasive, may not be detected using this algorithm. Early and advanced invasive squamous cancers are often erythematous and not aceto-white. We are continuing to explore the addition of fluorescent light to enable detection of those lesions that are not keratinized and thus not aceto-white.

A major advantage of digital colposcopy with automated image analysis is that it requires only limited resources, so clinical application of this technique could provide an inexpensive alternative to Pap screening and conventional colposcopic assessment. Furthermore, this technique could also provide a map showing where to perform biopsies for providers with less experience. Digital colposcopy, if the equipment can be produced inexpensively, may provide an alternative to DVI, VIA, and VILI for screening in low-resource settings. While DVI is less expensive in terms of instrumentation and methods, it is an entirely provider-dependent screening test. Its success depends on extensive formal training and experience. In the developing world, where health care resources are limited and there are few trained health care providers, screening tools with minimal training requirements are essential [32]. Digital colposcopy with automated image analysis has the potential to reduce the cost and time associated with training, leading to faster development and deployment of effective screening programs for cervical cancer throughout the developing world.

Chapter 5: Diagnostic Image Analysis Using Markov Random Field Model

5.1 INTRODUCTION

Cervical cancer is the second leading cause of cancer death in women worldwide and about a quarter million women die from cervical cancer each year [1]. This is mainly ascribed to the high mortality of cervical cancer in developing countries due to a lack of organized screening programs. However, cervical cancer is almost completely curable if it is detected early. In developed countries, effective screening programs to detect Human papillomavirus (HPV) infection and precancerous cervical lesions have significantly reduced the incidence and mortality of cervical cancer [1-4].

The main reason for the lack of healthcare screening programs in developing countries is a lack of sufficient resources. For example, the Papanicolaou smear is the most common screening method, but it requires significant laboratory infrastructure and highly trained personnel, which are not available in many developing countries [6-8].

As alternative screening techniques for low resource settings, VIA and DVI have been investigated by many researchers [7-9, 12, 14, 15]. Compared to the standard protocol of screening followed by diagnosis and treatment for cervical cancer patients, VIA or DVI simply allows a see-and-treatment approach. Large clinical trials of this technique are currently in progress in a variety of settings and results are becoming available to show the safety and efficacy of this approach [7-11].

Recently, digital optical imaging has emerged as another promising alternative to conventional screening methods. Optical imaging with automated image analysis provides the ability to quantitatively and non-invasively detect changes in cellular chemistry and in tissue architecture associated with progression of disease in real-time [17-23]. This approach also allows a see-and-treat procedure and, furthermore, it does not require extensively trained personnel since it does not depend on the provider's expertise to interpret an image as is the case for VIA or DVI.

Previous research has demonstrated that optical spectroscopy and imaging have the potential to improve screening and diagnosis of neoplasia [17-23, 26-40]. Huh et al. [23] measured the performance of optical detection of high-grade intraepithelial neoplasia (HGSIL) using fluorescence and reflectance spectroscopy, finding a sensitivity of 90% and a specificity of 70%. In a randomized clinical trial, Alvarez et al. [25] investigated the performance of an optical detection system that utilizes fluorescence, white light reflectance, and video imaging. This optical detection system, used as an adjunct to colposcopy, showed an increased ability to discriminate HGSIL from normal tissue compared to colposcopy alone.

Optical imaging systems and various image analysis techniques have been also studied in order to assess the entire field of the cervix for the purpose of locating the margins of neoplastic lesions [31, 33, 37, 39]. Li et al. [31] developed a computer-aided diagnosis system using colposcopic features such as acetowhitening changes, lesion margin and blood vessel structures. They designed image processing algorithms for anatomical features, acetowhite regions, and blood vessels. Balas [33] developed a multispectral imaging system which is capable of performing time-resolved multi-spectral imaging spectroscopy. The temporal changes in the intensity of backscattered light after application of acetic acid enabled differentiation between neoplastic and non-neoplastic lesions.

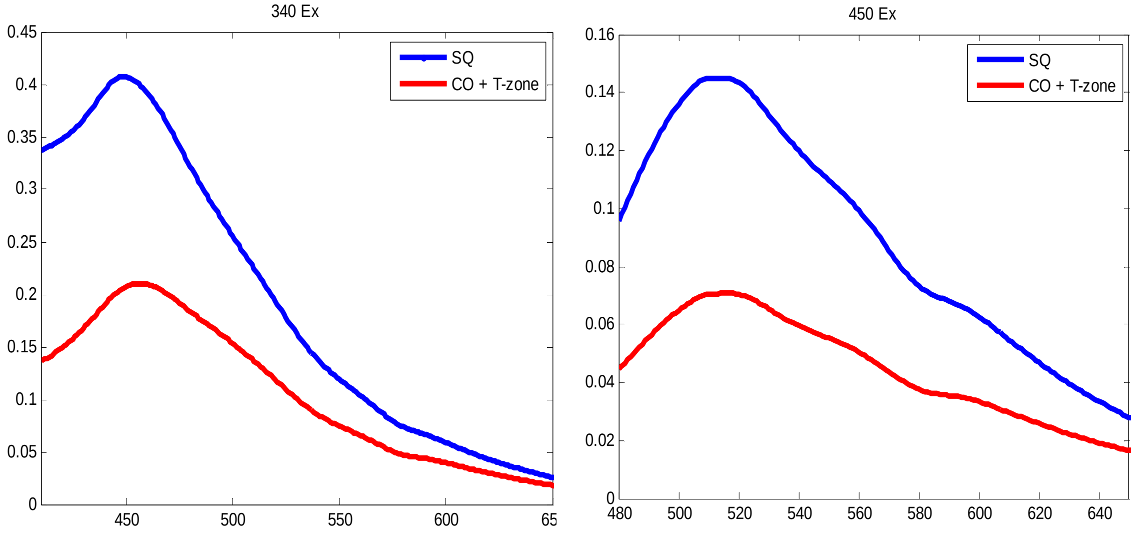
Even though the previous studies on optical imaging systems and image analysis techniques showed promising results, limitations remain. Current methods do not take advantage of the unique properties of cervical tissue directly in the design of image analysis algorithms. Instead, the image analysis techniques applied in most of the previous studies are generic and have not been tailored to utilize the unique optical features of cervical tissue. In this chapter, I explore the development of an image analysis framework which utilizes these features.

Cervical tissue can be mapped into regions characterized by a predominance of squamous tissue (ectocervix), columnar tissue (endocervix), and a transformation zone at the intersection between the two tissue types. Considering that cervical pre-cancer mainly arises in the cervical transformation zone, image analysis algorithms which initially identify tissue type may provide valuable diagnostic information. Furthermore, the optical properties of tissue in the cervix vary by tissue type, and, therefore, delineating cervical tissue types in an image creates an opportunity to improve the performance of a diagnostic system by accounting changes in baseline optical properties as a function of tissue type.

Recently, several other groups presented image analysis algorithms which consider tissue information. However, this work did not fully take advantage of this biological information in an integrated diagnostic image analysis framework. Gordon [34] applied an unsupervised Expectation-Maximization (EM) clustering algorithm to separate acetowhitening lesions from normal squamous epithelium and columnar epithelium using the color intensity of pixels and initial segmentation results were shown. However, no diagnostic framework was developed and the clinical performance using this technique was not reported. Gordon and Li [32, 36, 39] developed an image analysis algorithm to segment the anatomical features of cervix such as columnar epithelium, squamous epithelium, endo-cervical canal and the transformation zone based on color intensity values. This research also showed the potential of automatic segmentation of the transformation zone. However, they did not investigate any diagnostic features that consider tissue information, and, further, their analysis did not take into account the spatial relationship between tissue types. The diagnostic performance of this approach was not reported.

The purpose of the study in this chapter is to develop an integrated diagnostic image analysis framework using white light reflectance and fluorescence images. The focus of my approach is to exploit information specific to cervical tissue, at every stage of the proposed algorithm including classification as well as segmentation. Two clinical observations are used in this study. One is that most cervical precander or cancer occurs at the transformation zone [41], and the other is that the baseline optical properties of tissue vary as a function of tissue type. Figure 5.1 shows the average fluorescence spectrum measured from 82 patients using a FastEEM fluorescence/reflectance point probe system. In the figure, the blue and red curves correspond to the emission spectra of squamous epithelium and those of the transformation zone or columnar epithelium, respectively. For both 345 and 440 nm excitation, the emission spectra of squamous epithelium are uniformly higher in intensity than those of the transformation zone or columnar epithelium.

These two observations have important implications for optical imaging applications. Diagnostic image analysis algorithms should be designed differently according to the target tissue type, and the location of the tissue should be considered in analysis. In this study, I developed a probabilistic diagnosis framework to separate neoplastic from normal tissues using a Markov random field model. In order to consider detection of precancerous lesions within different tissue types, different features in squamous epithelium and in the transformation zone or columnar epithelium were explored. The spatial location of the tissue is considered in a Markov random field framework. We explore the potential of our diagnostic image analysis technique using white light reflectance and fluorescence images acquired from Multispectral Digital Colposcope (MDC).



(a) (b)

Figure 5. 1 Typical fluorescence spectra measured from normal squamous tissue (blue) and the transformation zone or columnar epithelial tissue (red) of cervix at 340nm (a) and 450nm (b) excitation.

5.2 MATERIALS AND METHODS

5.2.1 Instrumentation

The MDC was developed to acquire multispectral reflectance and fluorescence images of the entire cervix. The MDC consists of a commercially available, tilt-stand colposcope (Model 1DL, Leisegang, Germany) with a video rate color CCD camera and frame grabber for image acquisition. Details of this instrument can be found in papers by Benavides et al. [39] and Park et al. [70]. A Xenon arc lamp provides broadband illumination for reflectance imaging and monochromatic illumination for fluorescence imaging. Monochromatic light for fluorescence excitation is produced using 20 nm bandwidth band-pass filters centered at 345 nm and 440 nm excitation. The reflectance and the fluorescent light are captured and filtered using a commercially available video rate color charge-coupled device (CCD) camera (CV S3200 Rev B, JAI, Japan).

5.2.2 Clinical Study

The study protocol was reviewed and approved by the institutional review boards at The University of Texas M. D. Anderson Cancer Center, The University of Texas Health Science Center in Houston, The Lyndon B. Johnson County Hospital and the Harris County Health District, The British Columbia Cancer Agency, Rice University, and The University of Texas at Austin. Eligible patients were at least eighteen years old, not pregnant, and were referred to the colposcopy clinic with an abnormal Papanicolaou smear. Written informed consent was obtained from each participant. Following colposcopic examination, but prior to the treatment with a LEEP, white-light reflectance and fluorescence images excited at 345 nm and 440nm were acquired with the MDC from each patient at baseline. A second set of white-light reflectance and fluorescence images was acquired following the application of acetic acid (6%) on the cervix for two minutes. The patients then underwent a LEEP procedure. The specimen was inked and cut by the study pathologist. Each ectocervix was cut into twelve pieces. The study pathologists reviewed all the hematoxylin and eosin stained slides jointly, noting areas of cervical intraepithelial neoplasia (CIN). The histopathologic slides were scanned and the images were uniformly enlarged by 15% to account for shrinkage by formalin [71], under the assumption of uniform shrinkage. The images were then reconstructed into a three dimensional "pathology map" delineating areas of intraepithelial neoplasia using VolGen® (Austin, UT). Pathology maps were reviewed by the study clinicians and pathologists and compared to the white light reflectance and fluorescence images as shown in Figure 5.2.

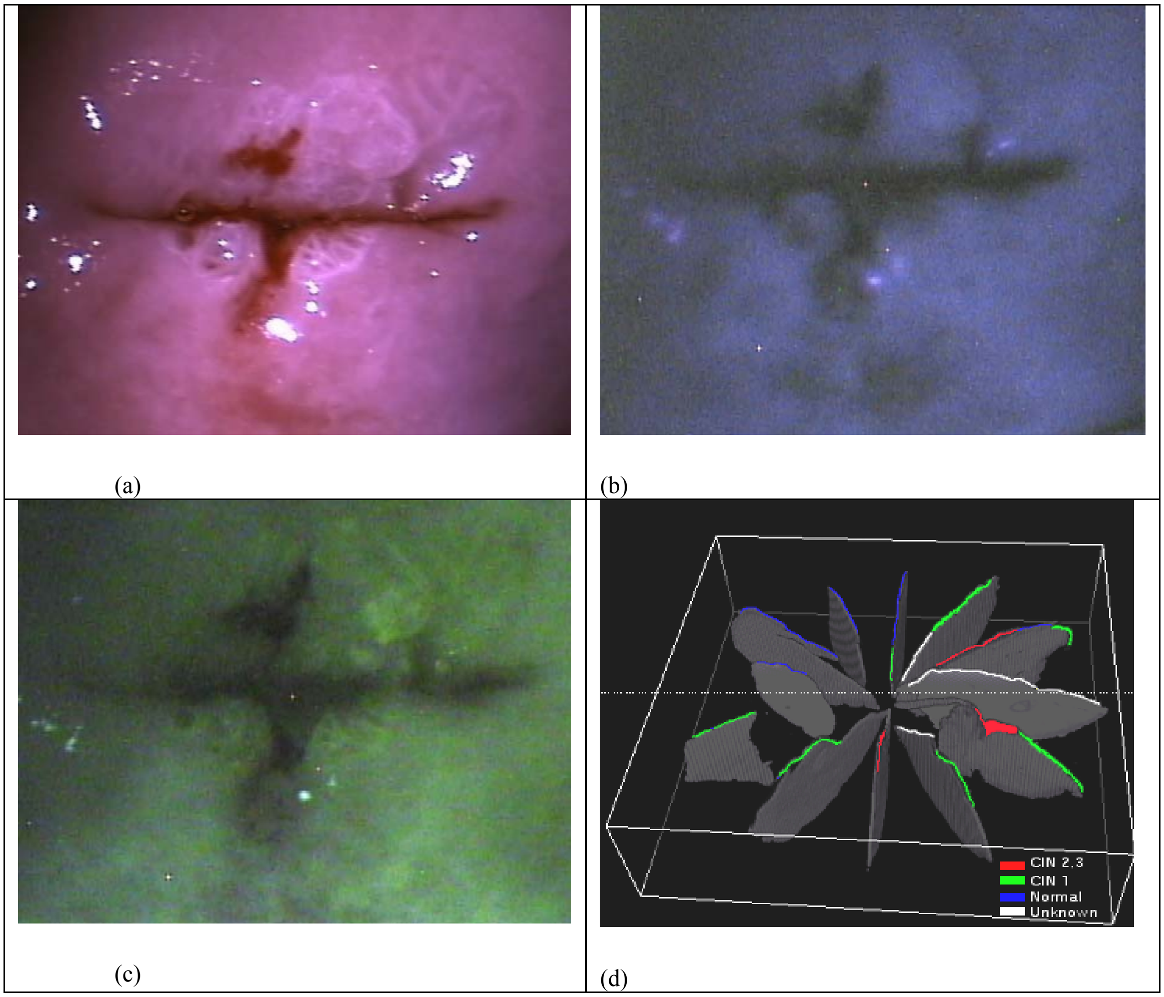

(a) (b) (c) (d)

Figure 5. 2 MDC image measurement and histopathology map. (a) White light reflectance image. (b) Fluorescence image at 345nm excitation. (c) Fluorescence image at 440nm excitation. (d) Histopathology map of the ectocervix.

5.2.3 Data Processing

The image analysis framework developed is depicted in Figure 5.3. The approach consists of three major steps which follow image preprocessing. The first step is a localized feature analysis algorithm developed separately for each tissue type. This algorithm was used to investigate which features are appropriate for separating precancerous or cancerous lesions from normal tissue in each tissue type. Secondly, tissue regions with similar optical properties for each tissue type were clustered. The final step is a classification algorithm using Markov Random Field model, which determines whether the clustered regions contain neoplastic lesions considering neighborhood information.

Image pre-processing

The first step for image pre-processing is to remove impulse noise spikes using 3x3 median filters. After noise removal, registration of images obtained before and after acetic acid application was performed. For the image registration, an automatic registration algorithm based on mutual information maximization [62, 63] was developed. In the registration algorithm, the location of the cervical os was used as a reference, and an automatic os detection algorithm was developed using the intensity and the location of the os.

In order to study the optical image properties of different tissue types, we developed a transformation zone detection algorithm based on pre acetic acid white light reflectance images. Pre acetic acid white light reflectance images were used because, compared to fluorescence images, they have better contrast between squamous epithelium and the transformation zone or columnar epithelium. Furthermore tissue information might become obscured after acetic acid application due to acetic acid changes induced in the appearance of neoplastic lesions. Based on the result of this algorithm, we also generated an epithelial map which indicates the location of the regions of squamous epithelium, columnar epithelium, os, any area obscured by specular reflection and image regions not occupied by tissue.

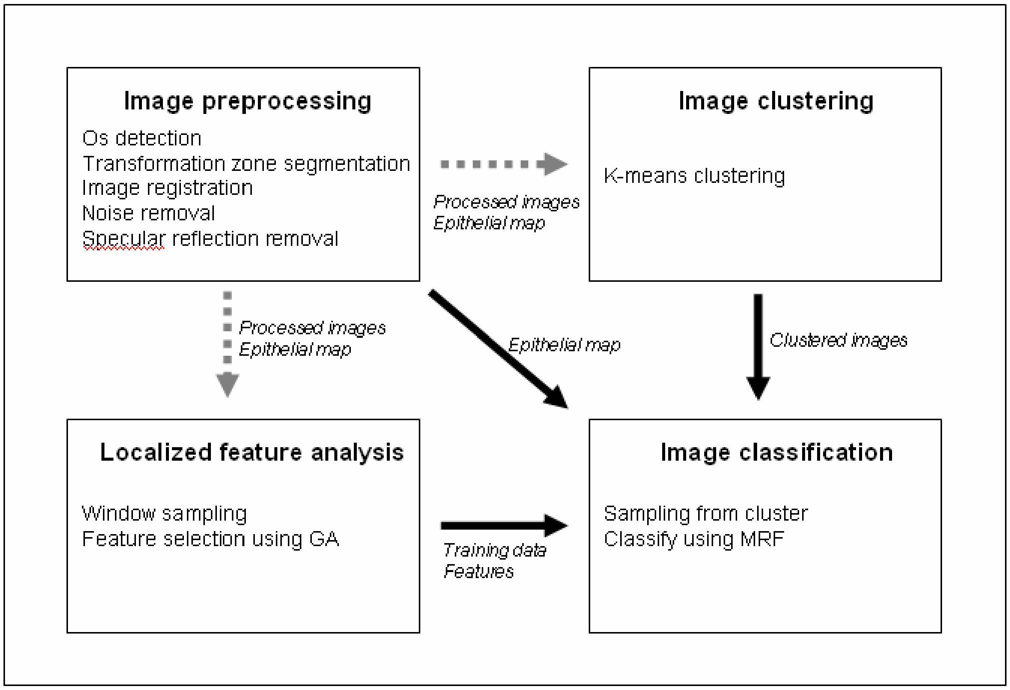


Figure 5. 3 A diagram illustrating the image analysis algorithm using tissue information. As the first step of the image analysis algorithm, image preprocessing was performed. Tissue regions with similar optical properties were clustered together. Finally, clustered images were classified based on optimal image features identified from the analysis of localized image regions and tissue type.

Localized feature analysis

To identify the most useful features to determine whether localized image areas contain neoplastic tissue, training data were obtained using a window-based sampling approach. 20x20 pixel windows were sampled from histopathologically proven areas of normal epithelium or low grade squamous intra-epithelial lesions (LGSILs) and high grade squamous intra-epithelial lesions (HGSILs). Since HGSIL is treated and LGSIL is followed, training data obtained from HGSILs were considered to be abnormal, while those from areas of normal tissue and LGSIL were considered to be normal.

For each window, five features including the mean, the standard deviation, the 90th percentile, the 95th percentile, and Shannon entropy were calculated from the original red-green-blue (RGB) color intensity and from ratios of individual RGB intensity of pre and post acetic acid white light reflectance and fluorescence images. This procedure was performed separately for each tissue type (squamous epithelium and columnar epithelium) using the epithelial map generated in the previous step.

Next, we performed feature selection to find diagnostically meaningful features which maximally separate normal and abnormal training data. Determination of the optimal set of features involves two individual problems: identifying both the optimal number of features and the optimal features themselves. In order to consider these problems simultaneously and to select the globally optimal feature set, a feature selection algorithm using genetic algorithms [77] was developed. The fitness function of the algorithm was defined using a measure considering inter- and intra-group Euclidean distances between normal and abnormal groups in the normalized feature space. The genetic operators were designed with stochastic parameters of 0.8 for the crossover probability and 0.1 for the mutation probabilities, which were determined in an empirical manner. In order to consider different tissue types, the proposed feature selection algorithm was applied to squamous epithelium and to columnar epithelium separately.

Using the selected features, we next performed a linear discriminant analysis per each tissue type. The classifiers were designed for different tissue types and were trained using leave-one-patient-out cross-validation. The algorithm is trained on all data excluding one patient and then applied to the held-out patient; then the process is repeated successively for all patients in the data set. The performance of classifiers was evaluated using receiver-operator characteristic curve (ROC) analysis. The classifiers separately trained for different tissue types were then used to classify entire images as described in the following sections.

Image clustering

Two image clustering algorithms were developed using the K-means clustering algorithm; results are compared to each other. The first algorithm considers only the pixel-wise intensities without considering tissue type, and the second utilizes information about tissue type as well. In the first algorithm, the whole image was clustered using the pixel-wise multi-modal intensity information including pre- and post-acetic acid image intensities both for white light reflectance and for fluorescence images. This algorithm utilized eight features from multi-modal intensities including intensity ratios of green to red, blue to red channels for pre- and post-acetic acid white light reflectance images, blue to red channels for pre- and post-acetic acid 345nm excited fluorescence images, and green to red channels for pre- and post-acetic acid 440nm excited fluorescence images. These eight features are summarized in table 5.1.

Table 5. 1 Intensity ratio features

Intensity ratios	White light reflectance images		345nm excited fluorescence images		440nm excited fluorescence images	
	Pre-acetic acid	Post-acetic acid	Pre-acetic acid	Post-acetic acid	Pre-acetic acid	Post-acetic acid
Green/Red	√	√	-	-	√	√
Blue/Red	√	√	√	√	-	-

In order to consider the relative contribution of these features for clustering, each of the features for clustering was weighted by the Shannon entropy $H(A)$ which is defined as [78]:

$$H(A) = -\sum p(a)\log_2 p(a),$$

where A is a random variable representing the value of a intensity feature, a is a realization of A ranging from 0 to 255, and $p(\bullet)$ is the probability mass (density) function of A.

We found that this approach, which clusters the image only based on image intensities, results in clusters that do not correspond well with the spatial anatomy of the images. In order to address this limitation, the second algorithm which utilizes tissue information was proposed. In this algorithm, we began with the epithelial map generated in the pre-processing step and clustered images depending on tissue type.

The performance of the two clustering algorithms was evaluated using the Davies-Bouldin (DB) index [79]. Within-cluster inertia W_k for the cluster k is defined as:

$$W_k = \frac{1}{n_k}\sum_{i \in C_k} d(x_i^k, x_c^k),$$

where x_c^k is the center of the cluster k, n_k is the number of points in cluster k, $d(x_i^k, x_c^k)$ represents a distance between two points x_i^k and x_c^k, and C_k is the set of points belonging to the cluster k. Between-cluster inertia B_{jk} for the cluster j and k is also defined as:

$$B_{jk} = d(x_c^k, x_c^j).$$

Then, the DB index λ^{DB} is defined by:

$$\lambda^{DB} = \frac{1}{K}\sum_{k=1}^{K} R_k,$$

where $R_k = \max_{j \neq k} \left\{ \frac{W_k + W_j}{B_{kj}} \right\}$ and K is the number of clusters. Using the DB index, the better performed clustering algorithm was identified by the higher DB index value.

Image classification using MRF

We developed a probabilistic image classification algorithm based on the Markov random field theory [80, 81]. The algorithm considers the neighborhood clusters' classification information in a probabilistic manner. Suppose that we have K clusters after clustering. Let y_i denote a set of the intensity data for pixels in cluster i. Using intensity data only, we can calculate the conditional probability distribution $p_{W_i|Y_i}(w_i \mid y_i)$ for each cluster $i \in \{1,...,K\}$, where $w_i \in \{0,1\}$ is the classification label vector for a cluster. Here we use the capitalized symbol as a random variable for realized data represented by the corresponding lower-case symbol, for example, X is the random variable for x.

In order to incorporate the neighborhood clusters' classification information, an augmented posterior probability distribution based on the Markov random field theory is proposed. Let ∂_i and w_{∂_i} denote the set of clusters neighboring cluster i and the collection of classification results of cluster i's neighboring clusters, respectively, that is, $w_{\partial_i} = \lfloor w_j \rfloor_{j \in \partial_i}$. Then the new posterior probability distribution $\widetilde{p}_{W_i|W_{\partial_i},Y_i}(w_i \mid w_{\partial_i}, y_i)$ is defined as:

$$\widetilde{p}_{W_i|W_{\partial_i},Y_i}(w_i \mid w_{\partial_i}, y_i) = \frac{1}{Z} e^{-\frac{U\left(w_i | w_{\partial_i}, y_i\right)}{T}},$$

where $U_i\left(w_i \mid w_{\partial_i}, y_i\right) = V_{i,1}\left(w_i \mid y_i\right) + V_{i,2}\left(w_i \mid w_{\partial_i}\right)$, $V_{i,1}\left(w_i \mid y_i\right) = -\ln p_{W_i|Y_i}\left(w_i \mid y_i\right)$, $V_{i,2}\left(w_i \mid w_{\partial_i}\right) = \sum_{j \in \partial_i} f_i(w_i, w_j)$, $f_i\left(w_i, w_j\right) = \begin{cases} \alpha_i & if\ w_i = w_j \\ \beta_i & if\ w_i \neq w_j \end{cases}$, and T is a parameter (temperature).

For classification, we use the maximum a posteriori (MAP) estimate, one of the well-known principles in Bayesian estimates. The MAP estimate w_i^{MAP} for W_i is obtained by:

$$w_i^{MAP} = \arg\max_{w_i} \tilde{p}_{W_i|W_{\partial_i},Y_i}\left(w_i \mid w_{\partial_i}^{MAP}, y_i\right),$$

where $w_{\partial_i}^{MAP} = \left\lfloor w_j^{MAP} \right\rfloor_{j \in \partial_i}$.

This equation is a standard formulation of MAP estimate; however, there is a complication for solving this equation. That is, since it involves MAP estimates of the neighboring clusters, it should be solved with the equations for the other clusters simultaneously.

To assess the classification results of each image, a grid-based approach for calculating sensitivity and specificity was used. In this approach, an image is divided into 20×20 grids and the classification results of the pixels in each grid of an image are compared with the corresponding pathology map. If the number of misclassified pixels is less than 20% of the number of pixels in the grid, then the grid point is considered to be correctly classified. Using this approach, the sensitivity, Se_{image}, and the specificity, Sp_{image}, of an image are defined as follows:

$$Se_{image} = \frac{number\ of\ grids\ correctly\ identified\ as\ positive\ grids}{number\ of\ positive\ grids},$$

$$Sp_{image} = \frac{number\ of\ grids\ correctly\ identified\ as\ negative\ grids}{number\ of\ negative\ grids}.$$

5.3 RESULTS

In this pilot study, pre- and post-acetic acid white light reflectance and fluorescence images were measured from twenty-one LEEP patients. Fifteen patients with CIN2 or 3 were considered to be abnormal patients while six patients with CIN1 or normal tissue were considered to be normal patients.

Figure 5.4 shows different optical properties of squamous epithelial tissue and columnar epithelial tissue in white light reflectance and fluorescence images. Figure 5.4(a) shows a patient's pre acetic acid white light reflectance image with the boundary of transformation zone enclosed with the red line. In this figure, the red color intensity of the transformation zone is higher than that of squamous epithelial tissue area. Vascular patterns of the transformation zone increase the red signal in the white light reflectance image. Corresponding fluorescence images excited at 345 nm (Figure 5.4(c)) and 440 nm (Figure 5.4(d)) show decreased fluorescence intensity in columnar tissue area. This finding is similar to the result from the point based spectroscopy study shown earlier in Figure 5.1. An automatic segmentation algorithm for the transformation zone and columnar epithelium was applied to the pre acetic acid white light reflectance image and the result is shown in Figure 5.4(b). Squamous epithelium, the transformation zone, os, background and specular reflection area are labeled with different colors. An H&E slide shown in Figure 5.4(e) was acquired from 9'o clock position of the patient's LEEP and shows clear morphological changes between squamous epithelium and columnar epithelium.

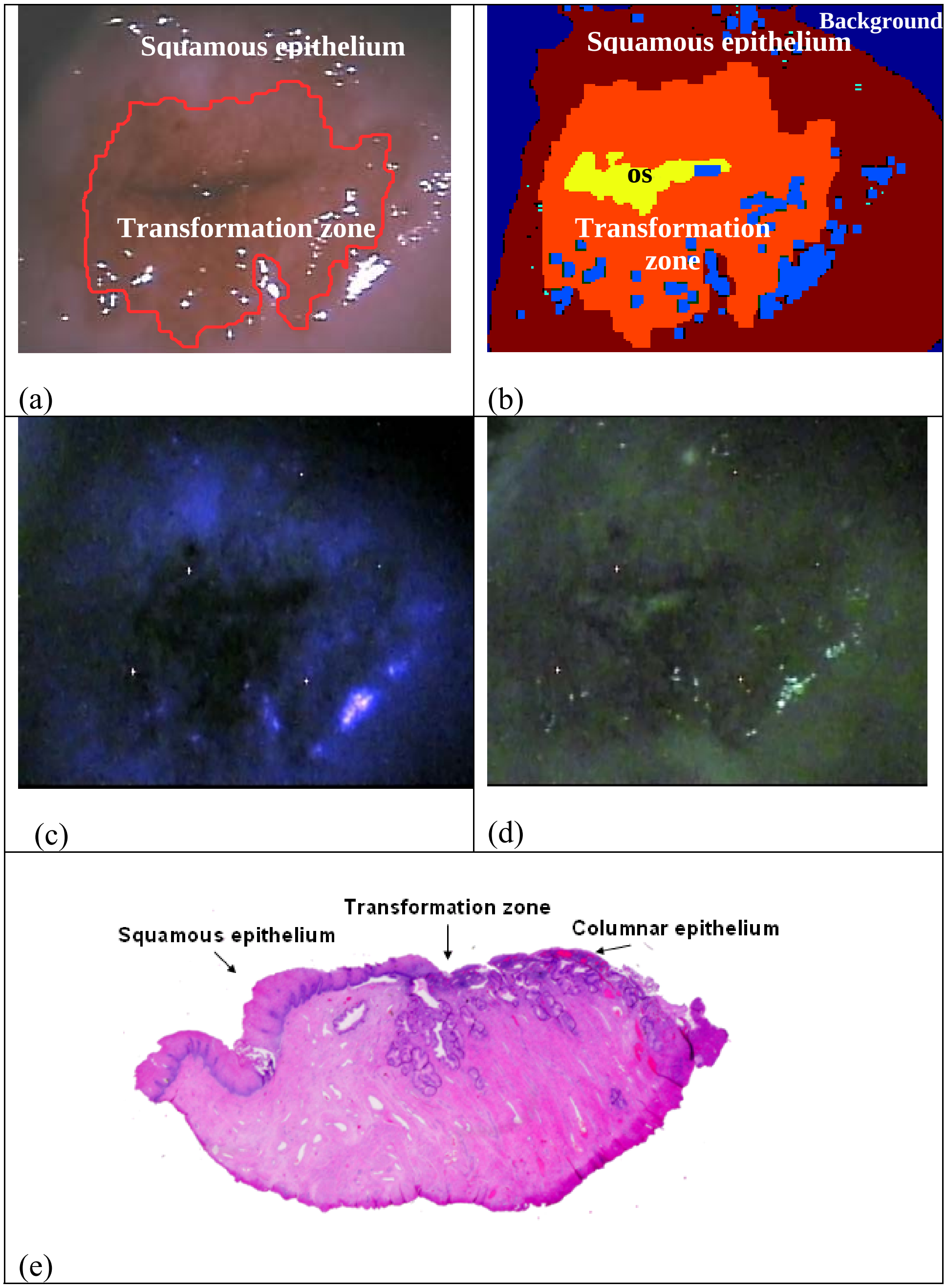


Figure 5. 4 Optical properties of different tissue types (squamous epithelium and transformation zone or columnar epithelium) (a) White light reflectance image (b) Epithelial tissue map (c) Fluorescence image at 345nm excitation. (d) Fluorescence image at 440nm excitation. (e) H&E slide

Figure 5.5 shows the clustering results of two patients. Figures 5.5(a) and (c) were clustering results using a K-means clustering algorithm without tissue information. Figures 5.5(b) and (d) were the results with tissue information. To compare two clustering algorithms' performance, inter-distance between clusters and intra-distance within each cluster were compared in Figure 5.5(e) and (f). We found that the clustering results using tissue information yield better performance. That is, the results using tissue information show smaller intra cluster distance and larger inter cluster distance.

Figures 5.6(a), (b), (c) and (d) show scatter plots of abnormal and normal samples from squamous epithelial tissue and the transformation zone or columnar tissue. The mean of the green to red intensity ratio versus the mean of the blue to red intensity ratio from white light reflectance images is shown in Figure 5.6(a) and (b). Noticeable separation between abnormal and normal samples can be observed for data sampled from squamous tissue. However, the same features from data from columnar tissue did not yield as clear a separation.

Results using the mean of green to red intensity ratio versus the standard deviation of green to red intensity ratio from fluorescence images excited at 440 nm are shown in Figure 5.6(c) and (d). Interestingly, the features from fluorescence images made the separation clearer between normal and abnormal data from the transformation zone or columnar epithelial area.

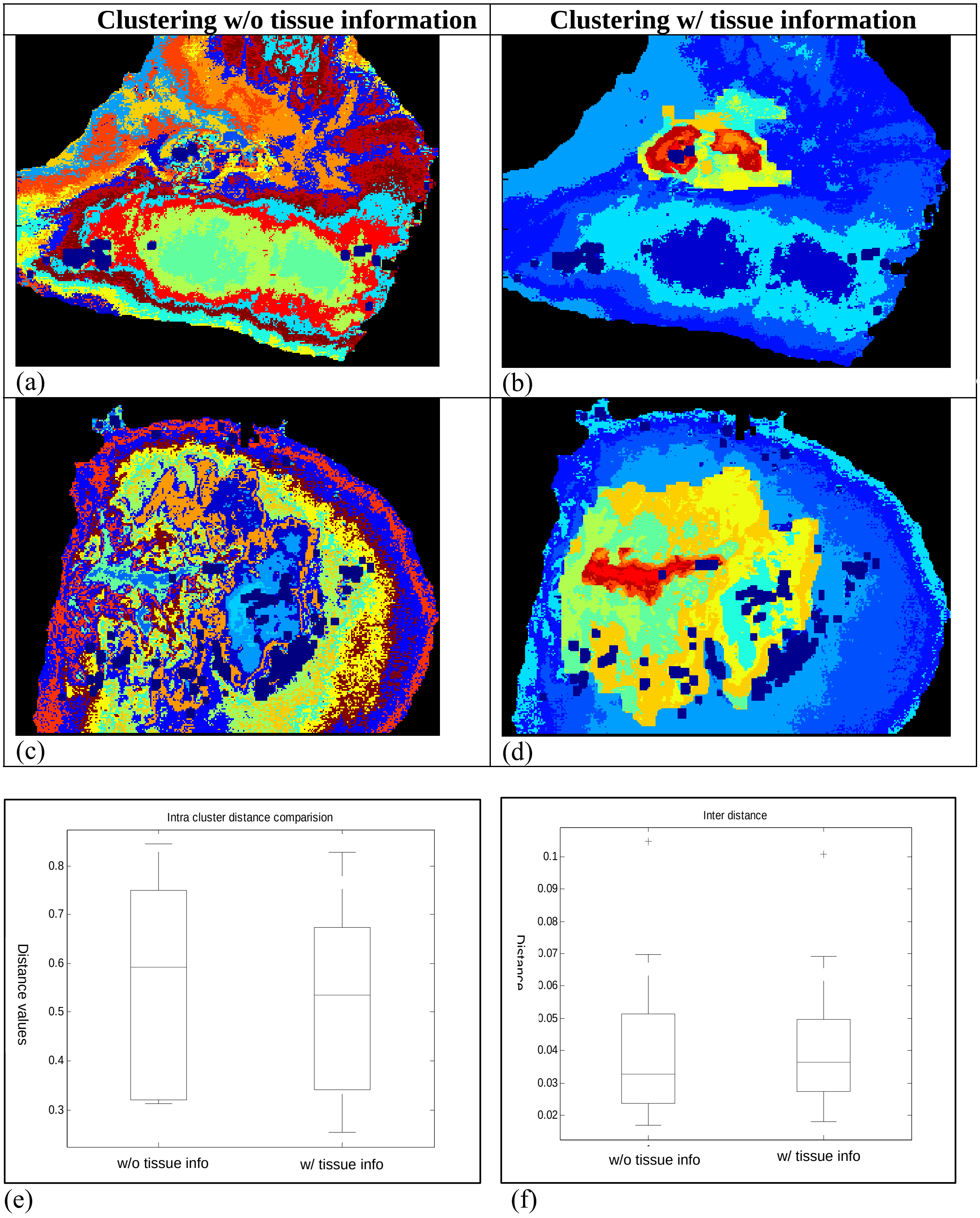


Figure 5. 5 Image Clustering Results without ((a) and (c)) and with ((c) and (d)) tissue information and comparison of intra- ((e)) and inter-distance ((f)).

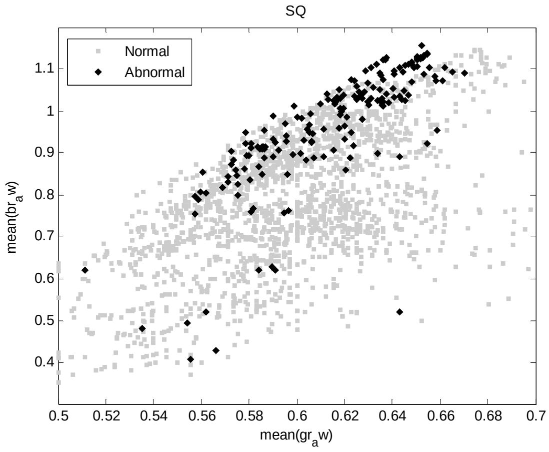
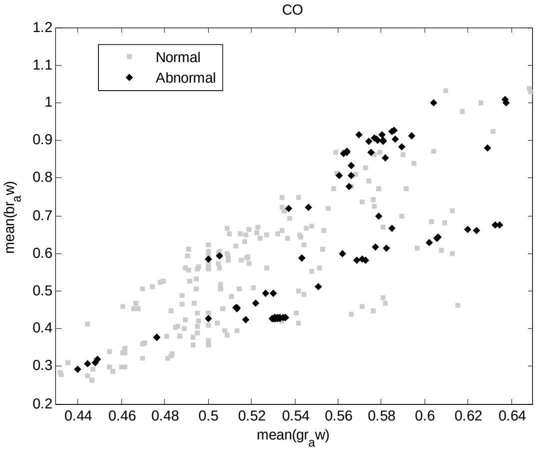
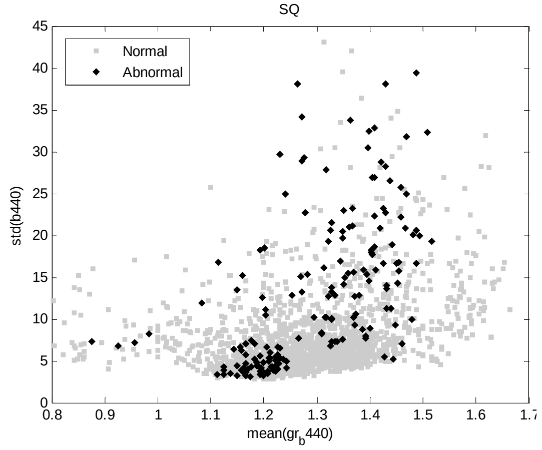
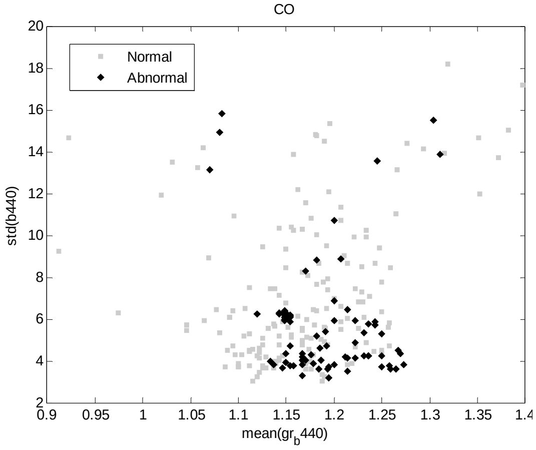
	Squamous epithelium	Transformation zone or columnar epithelium
Reflectance feature	 (a)	 (b)
Fluorescence feature	 (c)	 (d)

Figure 5. 6 Scatter plots of image features from normal and abnormal tissues. For reflectance images, the mean of green to red intensity ratio versus mean of blue to red intensity ratio is shown ((a) and (b)). For fluorescence images, the mean of green to red ratio versus standard deviation of blue channel for 440 excitation is shown ((c) and (d)). Left column plots ((a) and (c)) are for squamous epithelium and right column plots ((b) and (d)) are for transformation zone or columnar epithelium.

In order to compare classification performance for different tissue types, ROC curves were drawn in Figure 5.7 for data from squamous tissue, from columnar tissue, and from all tissue types. As shown in the figure, the data set from squamous tissue yielded the best classification result. The result for columnar tissue showed the least accurate performance. For each ROC curve, the point with the most desirable sensitivity and specificity values has been selected in Pareto-optimum sense [73], and the sensitivity and specificity values of the selected points are summarized in table 5.2.

Table 5. 2 Sensitivity and specificity values of Pareto-optimal points

	Squamous tissue	Columnar tissue	All tissue
Sensitivity (%)	80	60	80
Specificity (%)	79	70	75

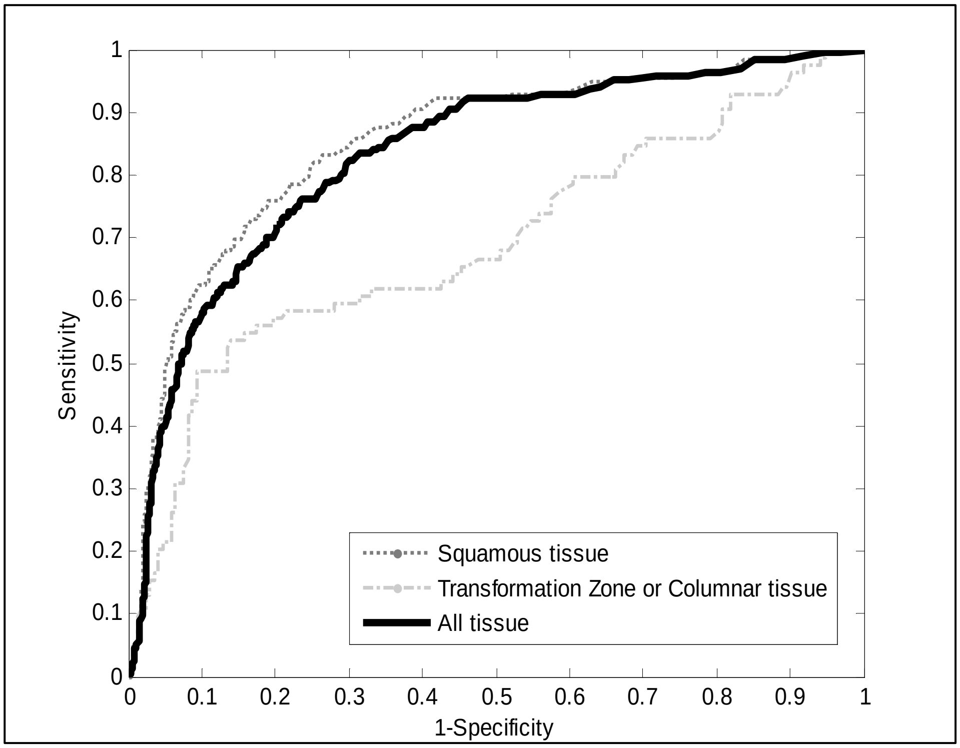


Figure 5. 7 ROC (Receiver Operating Characteristic) curves for classification between high grade SIL and normal or low grade SIL samples from squamous tissue, from the transmission zone or columnar tissue, and from all tissue types.

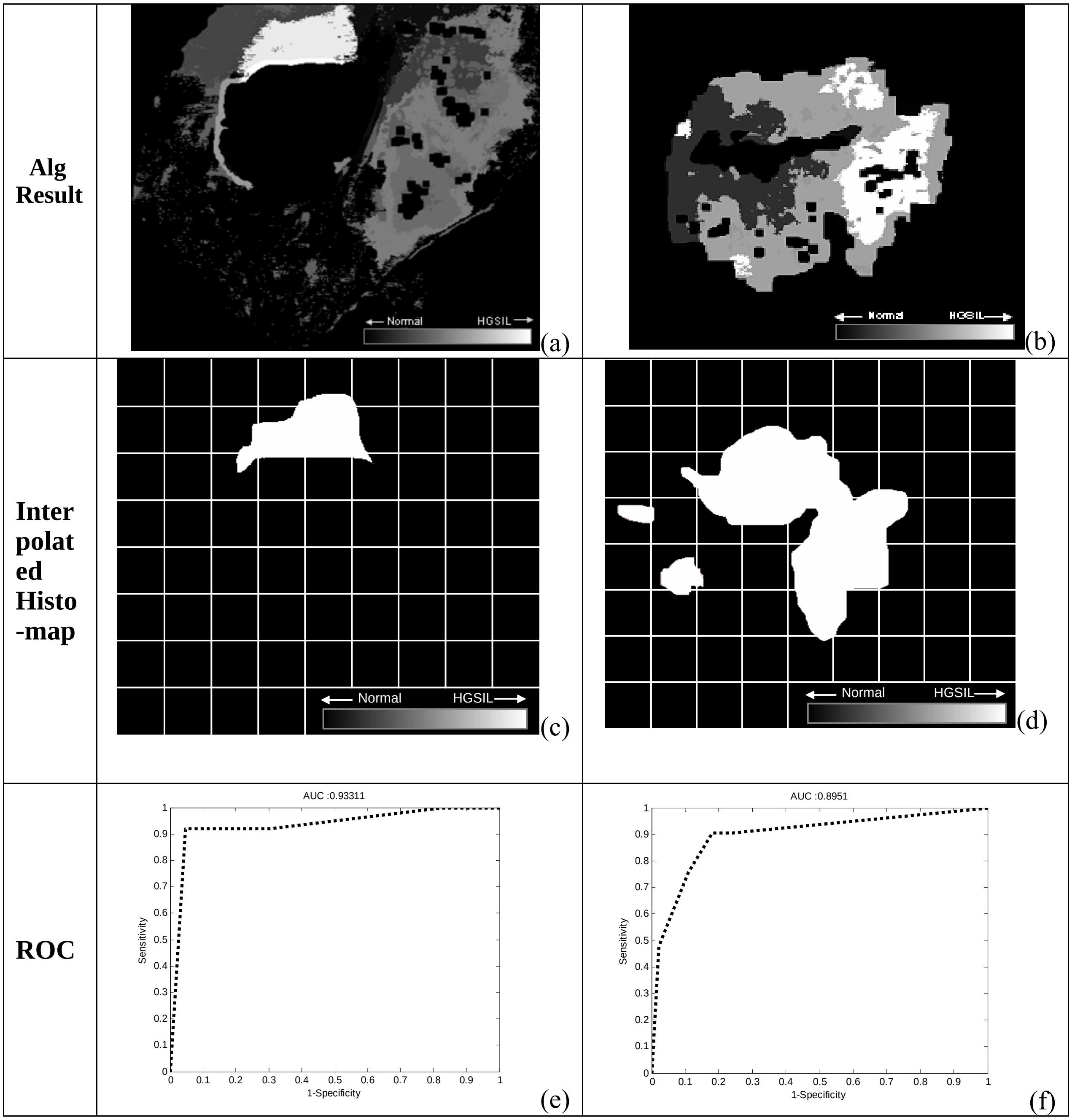


Figure 5. 8 Image classification results for two patients with CIN 2 and 3 lesions using the proposed image analysis algorithm. First row ((a) and (b)) shows algorithm results, second row ((c) and (d)) shows interpolated histo-maps, and third row ((e) and (f)) shows ROC curves assessed using the grid-based approach.

Table 5. 3 Performance of the image analysis algorithm

Patient No	Diagnosis of ecto cervix	AUC	Sensitivity	Specificity
1	CIN 2	0.96	92%	95%
2	CIN 2/3	0.78	65%	79%
3	CIN 1	-	-	100% (0.5 threshold)
4	CIN 2	0.5	50%	50%
5	CIN 2	0.99	100%	95%
6	CIN 1	-	-	100% (0.5 threshold)
7	CIN 2/3	0.90	89%	82%
8	CIN 2/3	0.5	50%	50%
9	CIN 2/3	0.76	62%	72%
10	CIN 2/3	0.56	50%	56%
11	CIN 2	0.93	84%	91%
12	CIN 2/3	0.84	70%	90%
13	CIN 2/3	0.75	60%	80%
14	CIN 2	0.82	65%	87%
15	CIN 2	0.54	51%	50%
16	Normal	-	-	100% (0.5 threshold)
17	Normal	-	-	100%
18	CIN 2	0.81	65%	93%
19	CIN 2	0.97	90%	89%
20	Normal	-	-	98% (0.5 threshold)
21	Normal	-	-	100% (0.5 threshold)
Average	-	0.77	70%	84%

Figure 5.8(a) and (b) show image classification results using the proposed algorithm based on the Markov Random Field model. The corresponding interpolated histopathology maps Figure 5.8(c) and (d) were generated from regions where the histopathology map indicated high grade disease following the guidance of the colposcopic impression provided by the physician. The impression is recorded by

marking on the white light reflectance image. From the comparison of each pair of classification results and the interpolated histopathology map, we can see that the high-probability regions (brighter area) in the disease probability map correlate well with regions of HGSIL in the interpolated histopathology map. The performance of the classification for each patient's data was assessed using a grid based approach, and the resulting ROC curves are shown in Figure 5.8(e) and (f).

Table 5.3 summarizes image classification results for 21 patients. In this table, diagnosis result, area under the ROC curve (AUC), sensitivity, and specificity of each patient are listed. For patients diagnosed either as normal or as CIN 1, ROC curves could not be drawn, and only specificity values were calculated using the threshold probability of 0.5. Furthermore, in the last row of the table, the average values of AUC, sensitivity and specificity are specified for all patients in the study.

5.4 DISCUSSION

In this study, we proposed a novel automatic margin detection algorithm for the detection of cervical cancerous or precancerous lesions using multispectral digital colposcope images. Our results show the diagnostic potential of our technique with an average sensitivity of 70% and a specificity of 84% for detection of neoplastic areas from twenty-one patients.

From our image analysis, we have obtained two interesting findings. First, as we incorporated tissue information in our clustering algorithm, the spatial information helped the algorithm to generate better localized clusters with larger inter and smaller intra distances than those from algorithms which did not include this information. The better localized clusters resulted in better diagnostic results in this image classification stage.

The second finding is that features selected from reflectance or from fluorescence images show selective performance according to different tissue types. Our analysis results suggest that features selected from white light reflectance images show better performance for precancerous lesion detection in squamous tissue area. On the other hand, in columnar tissue area, the diagnostic performance of features from fluorescence images is superior to that of features from white light reflectance images. From this observation, we believe that incorporating tissue types to a multispectral digital colposcopic image analysis tool is crucial to maximize diagnostic performance. However, this finding must be explored more fully in the context of a larger clinical data set.

A potential weakness of our study is that our approach is limited to the detection of neoplastic lesions within the ectocervix only. In order to find lesions located on the endocervix, additional information is required. Additionally, specular reflection present sometimes in white light reflectance images may obscure the diagnostic signal from neoplastic lesions located in that area. Since the proposed MRF model utilized neighbor tissues' diagnosis results, tissues neighboring the specular reflection areas also can not be diagnosed properly. This limitation may deteriorate the diagnostic performance of our method. In chapter 6, we consider cross polarized imaging to remove the specular reflection following the approach by Li [40].

A major strength of the proposed approach is that it utilizes information about cervical anatomy and its relationship to cervical cancer progression in an image analysis model. We have explored the use of tissue type in the clustering and classification algorithms in our MRF model in order to consider spatial information in a probabilistic sense. Our results indicate that incorporating biological properties specific to the target clinical application into diagnostic image analysis algorithms can significantly improve the diagnostic performance of image analysis algorithms.

Our approach allows a provider to quantitatively asses the margin of neoplastic lesions. In the clinic, this technique may provide guidance to a provider to perform a LEEP or point spectroscopy procedure in an objective manner. If the proposed approach can be further automated, it may ultimately provide a more cost effective screening solution in low-resource settings by reducing the need for extensive provider training and expertise.

Chapter 6: Cross Polarized Image Analysis for the Detection of Cervical Cancer

6.1 INTRODUCTION

Cervical cancer comprises approximately 12% of all cases of cancers in women worldwide [1] and, in developing countries, it is a leading cause of cancer-related death. Furthermore, it is estimated that as many as one million new cases of cervical cancer could develop by 2050 [1]. Yet, cervical cancer is preventable through early detection. In developed countries, common screening programs including the Papanicolaou (Pap) smear have been effective in reducing the incidence and mortality of cervical cancer [2-4]. However, in developing countries, national cervical cancer prevention programs are rarely implemented since most of the current screening techniques reply on sophisticated laboratory procedures performed by highly trained personnel and many developing countries have only limited resources [1-4].

In developing countries, simple visual screening methods such as visual inspection with acetic acid (VIA) or direct visual inspection (DVI) have been introduced as low-cost alternatives to the Pap smear [5-8]. DVI or VIA are both techniques that allow a provider to visualize the cervix with either the naked eye or with low power magnification after the application of acetic acid [5, 7]. These techniques do not require a laboratory infrastructure since they utilize see-and-treat strategies. However, the diagnostic performance of these methods highly depends on provider expertise [6, 7] and, therefore, the diagnostic results can be inconsistent and hard to reproduce. In order to reduce this risk, it is crucial to develop objective criteria for the identification of positive lesions and to train personnel to correctly implement these criteria are crucial.

Recently, digital colposcopy with automated image analysis techniques has been investigated as an alternative screening method to the conventional techniques used in developing countries [26-40]. Several studies have shown the potential to discriminate an area of neoplastic tissue from normal tissue using diagnostic image analysis tools [30-40]. Li introduced a computer aided diagnosis system for cervical cancer using diagnostically important features such as acetowhitening, blood vessel structures and lesion margins [31, 38]. In chapter 5, a diagnostic performance of a digital image analysis tool using tissue information and the acetowhitening feature is introduced. In this study, the proposed diagnostic algorithm utilized tissue type information as well as neighborhood diagnosis information, yielding a diagnostic algorithm tailored to detection of cervical cancer. This pilot study showed a relatively high diagnostic performance. However, this study also illustrated a number of important limitations. First, unpolarized white light imaging prevented obtaining diagnostic information from areas obscured by specular reflection. Since the proposed Markov random field (MRF) model in this study utilized diagnostic information from surrounding tissue areas, the effect of specular reflections extended into neighboring regions. Second, better tissue type separation may improve diagnostic performance since the study utilizes neighborhood tissue diagnosis information.

Following the approach presented in chapter 5, this study proposes a diagnostic algorithm for cervical cancer utilizing tissue type information. Unlike this approach, we propose utilizing a different modality of measurement, that is, cross polarized imaging.

By adopting cross polarized imaging, we reduce the effect of specular reflection and we enhance the ability to segment images based on tissue type which may both lead to better diagnostic performance. In order to substantiate our expectation, we apply this approach to images acquired using various imaging techniques including unpolarized reflectance imaging, and cross- and parallel-polarization imaging. We quantitatively compare the diagnostic performance achieved using each of these imaging techniques.

6.2 MATERIALS AND METHODS

6.2.1 Instrumentation

In order to image the entire cervix, a multispectral digital colposcope (MDC) was developed to acquire multispectral reflectance and fluorescence images [39, 70]. The MDC system is composed of a light source, standard colposcope and a digital color charge-coupled device (CCD) color camera. Details of this instrument can be found in chapter 2.

6.2.2 Clinical Study

The study protocol was reviewed and approved by the institutional review boards at The University of Texas M. D. Anderson Cancer Center, The University of Texas Health Science Center in Houston, The Lyndon B. Johnson County Hospital and the Harris County Health District, The British Columbia Cancer Agency, Rice University, and The University of Texas at Austin. Patients 18 years and older who were not pregnant and who were referred to the colposcopy clinic with an abnormal Papanicolaou smear were eligible for this study. Written informed consent was obtained from each participant. Following colposcopic examination, but prior to the treatment with a LEEP, background, unpolarized white-light reflectance, cross and parallel polarized images were acquired with the MDC from each patient at baseline. A second set of images was acquired following the application of acetic acid (6%) on the cervix for two minutes. The patients then underwent a LEEP procedure. After the LEEP, in order for histopathology map construction, the margin boundary of LEEP was carefully marked on one of the acquired MDC images. The specimen was inked and cut by the study pathologist. At each trial site, pathologists reviewed all the hematoxylin and eosin stained slides jointly, noting areas of cervical intraepithelial neoplasia (CIN), squamous epithelium, columnar epithelium, and transformation zone. The histopathologic slides were scanned preserving real tissue size and then the images were uniformly enlarged by 15% to account for shrinkage by formalin [71], under the assumption of uniform shrinkage. The histopathology images were then reconstructed into a three dimensional "pathology map" delineating areas of intraepithelial neoplasia using VolGen® (Austin, Texas). When the map was constructed, the ratio between the actual size of each pixel of MDC-measured images and that of the scanned images was calculated to establish the appropriate mapping between the MDC image and the histopathology map. The position of the LEEP specimen boundary and the transformation zone were also incorporated to correlate between MDC images and the histopathology images. After construction, pathology maps were reviewed by the study clinicians and pathologists and compared to the reflectance images.

6.2.3 Data Processing

This section explains the main algorithm for data processing in order to compare the classification performance of images acquired under three different settings such as unpolarized white light, cross- and parallel-polarization. The classification algorithm is essentially the same as that described in chapter 5 and each of the major steps in the data processing is explained in the following sections.

A. Preprocessing

The first step for image pre-processing is to remove impulse noise spikes using 3x3 median filters. The noise removal procedure was performed for each of the background, cross-polarized, parallel-polarized, and unpolarized images. Then, the background image intensity was subtracted from that of each of the other images. After background subtraction, os detection and epithelial layer segmentation algorithms were applied to each of the images in the same manner described in chapter 5. Utilizing the detected os, registration of the images was performed using the automatic registration algorithm developed in chapter 5.

B. Localized feature analysis

Following the procedure described in chapter 5, localized feature analysis was performed to identify diagnostically useful features. The analysis was individually performed for each of three measurement modalities, i.e., unpolarized white light and cross- and parallel-polarized imaging. In addition to three measurements, we analyzed a derived image $\mathrm{Pol} = (I_{\mathrm{parallel}} - I_{\mathrm{cross}})/(I_{\mathrm{parallel}} + I_{\mathrm{cross}})$, where I_{parallel} and I_{cross} are parallel and cross polarized image intensity, respectively, following Jacques' method [66]. Training data were individually obtained for each imaging technique using a window-based sampling approach.

The five features, including the mean, the standard deviation, the 90th percentile, the 95th percentile, and entropy for individual values and for the ratios of the red-green-blue (RGB) color intensity normalized by the sum of three colors' intensities, were obtained for each tissue type. Table 6.1 shows various image intensity ratios from which each set of five features were calculated.

Table 6. 1 Intensity ratios for features

Intensity ratios	White light reflectance images		Cross polarized images		Parallel polarized images		$(I_{\text{parallel}} - I_{\text{cross}})/(I_{\text{parallel}} + I_{\text{cross}})$	
	Pre-acetic acid	Post-acetic acid	Pre-acetic acid	Pre-acetic acid	Post-acetic acid	Post-acetic acid	Pre-acetic acid	Post-acetic acid
R/(R+G+B)	√	√	√	√	√	√	√	√
G/(R+G+B)	√	√	√	√	√	√	√	√
B/(R+G+ B)	√	√	√	√	√	√	√	√
G/R	√	√	√	√	√	√	√	√
B/R	√	√	√	√	√	√	√	√

Using the features obtained in this manner, we applied a feature selection algorithm based on genetic algorithms [77] to find diagnostically meaningful features that maximally separate normal and abnormal training data. In this study, HGSILs were considered to be abnormal, while areas of normal tissue and LGSIL were considered to be normal.

Finally, a linear discriminant classifier was developed using leave-one-patient-out cross-validation training. The classifier performance was individually evaluated using Receiver Operator Characteristic curve (ROC) analysis for each of unpolarized white light, cross polarized, and parallel polarized imaging.

C. Image clustering and classification

Following the approach described in chapter 5, we first performed image clustering and, then, classified each cluster using the clustering results. For image clustering, the algorithm described in chapter 5, which utilized tissue type information, was adopted. After clustering, a probabilistic image classification algorithm was applied. The classification algorithm follows the algorithm described in chapter 5, and the maximum a posteriori (MAP) estimate was used based on Markov random field (MRF) theory. A grid-based approach for calculating sensitivity and specificity was also used to assess the classification results of each image as described in chapter 5.

6.3 RESULTS

Pre- and post-acetic acid images were measured using unpolarized white light, cross polarized, and parallel polarized reflectance imaging from twenty-one LEEP patients from M. D. Anderson (Houston, US), the British Columbia Cancer Agency (Vancouver, CA), and University College Hospital (Ibadan, Nigeria). Our site pathologists confirmed fourteen patients (67%) with CIN 2 or 3 lesions in the ectocervix, two patients (9%) with CIN 1 lesions, and five patients (24%) with no evidence of CIN in the ectocervix.

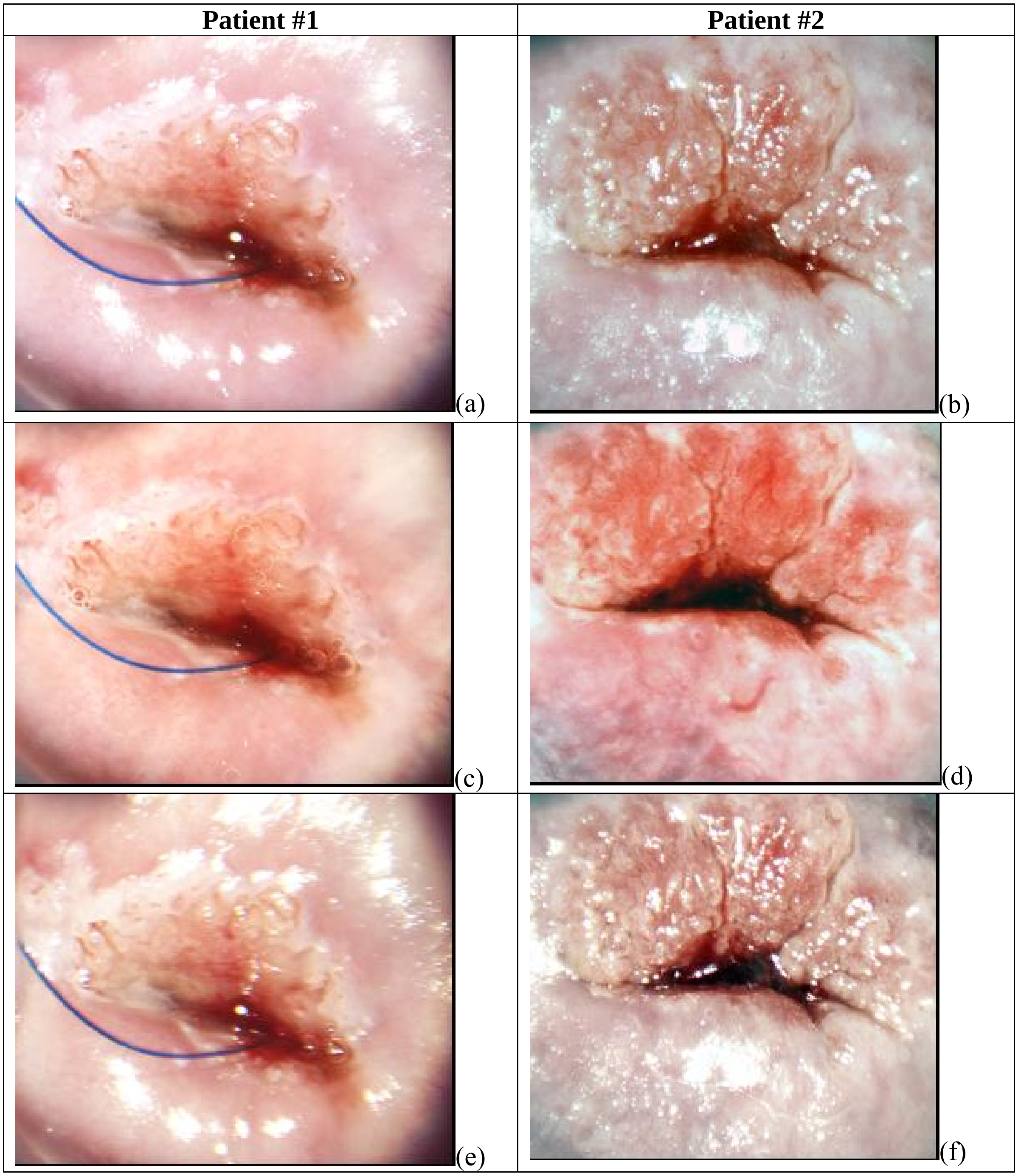
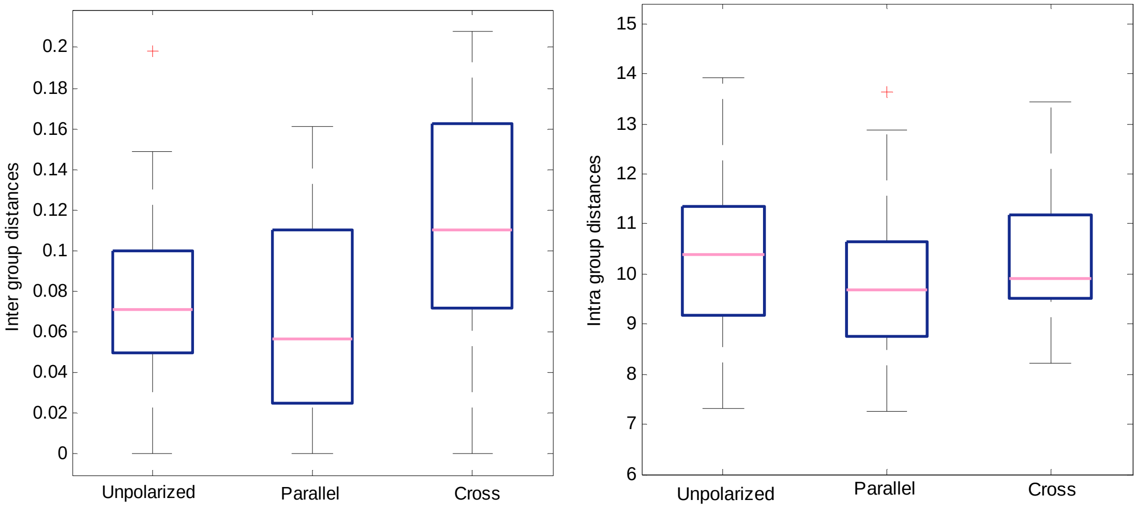


Figure 6. 1 MDC images (displayed on the previous page; (a) and (b) were measured using unpolarized imaging, (c) and (d) using cross polarized imaging, and (e) and (f) using parallel polarized imaging.)

Figure 6.1 shows typical post acetic acid reflectance images measured by the MDC from two patients in the study. Figure 6.1(a) and (b) in the first row show unpolarized reflectance images; significant specular reflections are observed. The second row (Figure 6.1(c) and (d)) shows cross polarized images. Specular reflections were clearly removed. Vascular structures were also more vividly shown in the cross polarized images. Figure 6.1(e) and (f) show parallel polarized images.



(a) (b)

Figure 6. 2 Inter group ((a)) and Intra group ((b)) distances between Squamous and Columnar or the transformation zone tissue

Figure 6.2 illustrates the inter- and intra- group distances between squamous tissue and columnar tissue acquired from unpolarized, parallel polarized, and cross polarized images. Squamous tissue were sampled from pixels in areas of squamous epithelium, and columnar tissue from pixels in areas of columnar epithelium or the transformation zone. As illustrated in the box plots in figure 6.2(a), the inter group

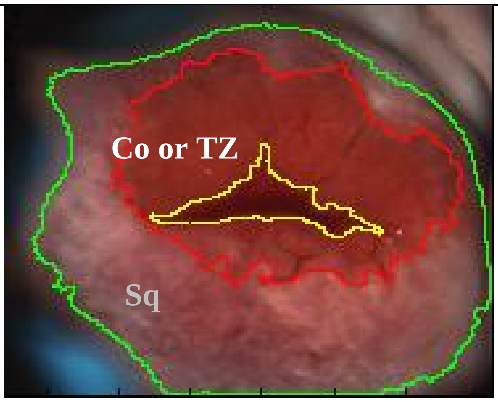
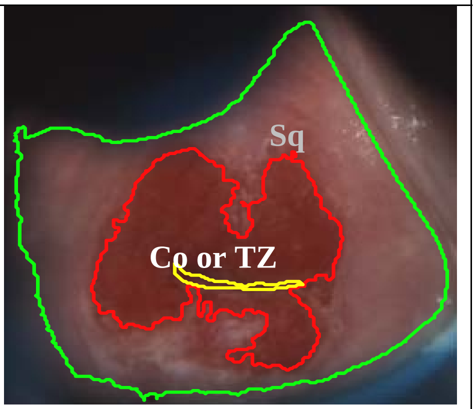
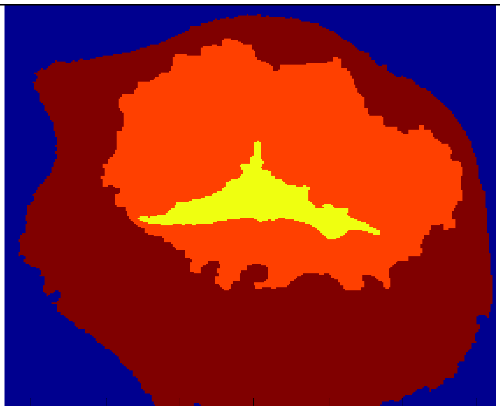
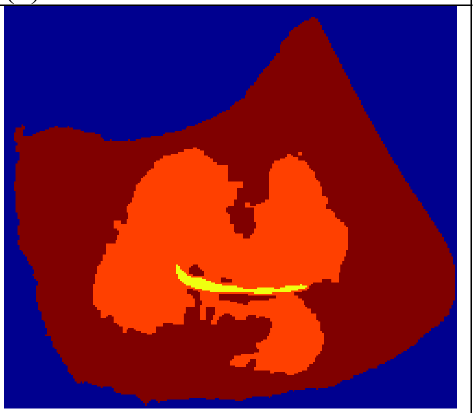
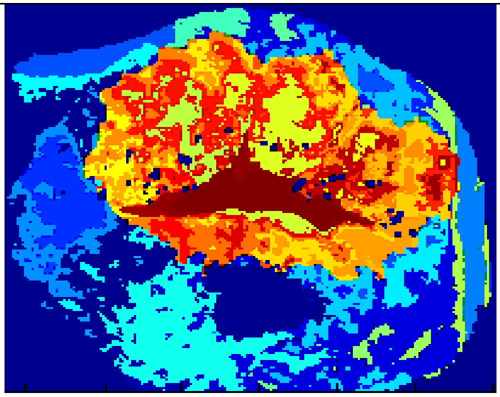
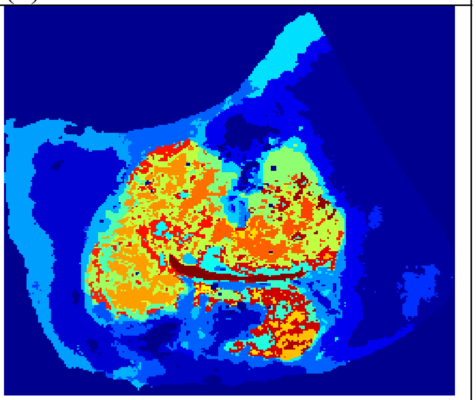
	Patient No. 1	**Patient No. 2**
Cross-polarized image	 (a)	 (b)
Epithelial map	 (c)	 (d)
Clustering result	 (e)	 (f)

Figure 6. 3 Os and Transformation zone detection, and clustering results using cross polarized images. In the figure (a) and (d), Co stands for Columnar tissue, TZ for transformation zone, and Sq for squamous tissue.

distance between squamous tissue and columnar tissue acquired from cross polarized images is bigger than that from unpolarized or parallel polarized images. Intra group distances between squamous data and columnar data do not show significant differences among unpolarized, parallel polarized, and cross polarized images.

Figure 6.3 shows the results of image clustering using the cross polarized images. In this clustering, epithelial maps which separate tissue types were best generated using pre acetic acid cross polarized images. Then, a K-means clustering algorithm was applied using tissue information acquired from epithelial map. In the figure, each row presents a cross polarized image (Figure 6.3(a) and (b)), the epithelial map (Figure 6.3(c) and (d)), and the clustering result (Figure 6.3(e) and (f)) for each patient.

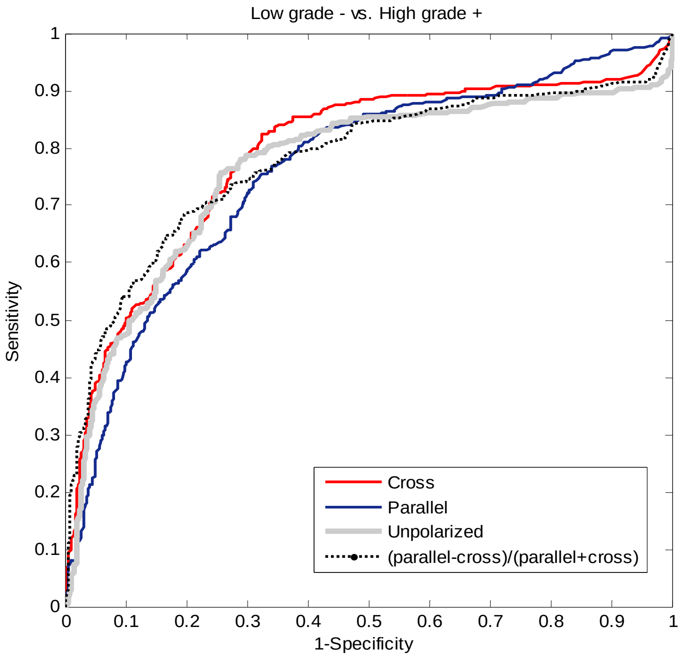


Figure 6. 4 ROC (Receiver Operating Characteristic) curves for the classification of high grade SIL vs. normal or low grade SIL samples from unpolarized, cross polarized, parallel polarized, and polarization ratio images.

The diagnostic performance of the algorithms based on localized features has been assessed for each of four training data groups with different data sources of unpolarized, cross polarized, parallel polarized, and polarization ratio images. Figure 6.4 shows a comparison of the results of performance assessment using ROC curves. Area under the curve (AUC) values of training data acquired from cross polarized images, from unpolarized images, from parallel polarized images, and from polarization ratio are 0.77, 0.74, 0.74, and 0.75, respectively. This observation suggests that the classification result for cross polarized images is similar with those for the other types of images.

Figure 6.5(a) and (b) show image classification results overlaid on cross polarization images. High disease probability regions identified from the classification were overlaid with red color. The corresponding histopathology maps (Figure 6.5(c) and (d)) were generated based on the information about regions where the histopathology map indicated high grade disease following the guidance of the colposcopic impression provided by the physician. The impression was recorded by marking on the cross polarized image immediately following LEEP. By comparing the classification result and the interpolated histopathology map, we find that the high-probability regions (dark red area) correlate well with regions of HGSIL in the interpolated histopathology map. In a more quantitative manner, the performance of the classification algorithm for each patient's data was assessed using a grid based approach, and ROC curves were drawn as shown in Figure 6.5(e) and (f). The Image classification result using cross polarized images is almost uniformly superior to that using parallel polarized or unpolarized images.

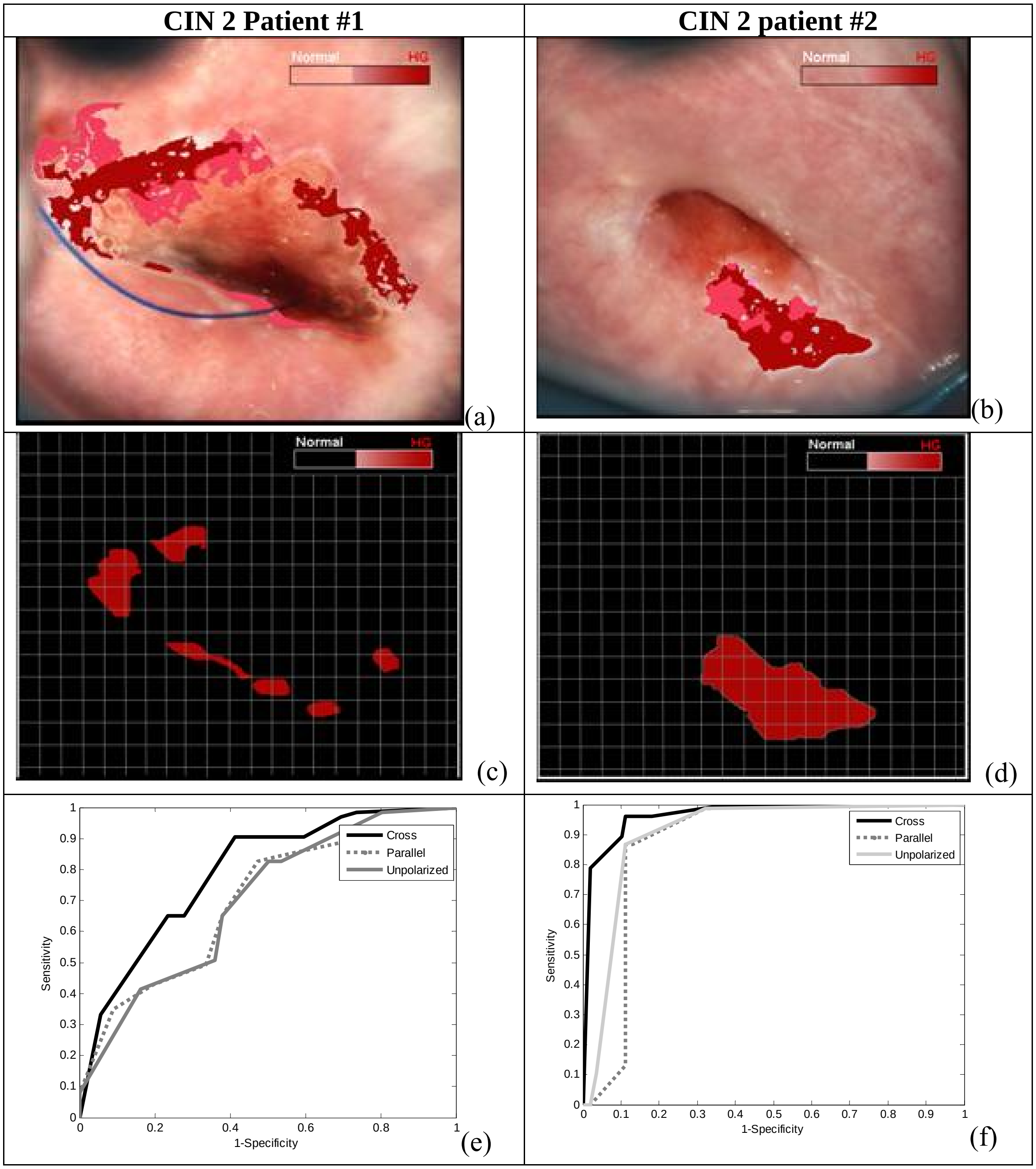


Figure 6. 5 Image classification results for two patients with CIN2. Each column shows three figures for each patient including a cross polarized image ((a) and (b)), a classification result using cross polarized images ((c) and (d)), and the ROC curves for cross polarized, unpolarized, and parallel polarized imaging ((e) and (f)).

Figure 6.6 shows the ROC curves for the twenty-one patients in this study as well as the average performance using ROC curves. The cross polarized image classification result for each patient was assessed using a grid-based approach. We have selected the 'desirable' point in the ROC curve of average classification performance in a Pareto optimal sense [73] and the corresponding sensitivity and specificity values are 70% and 71%, respectively.

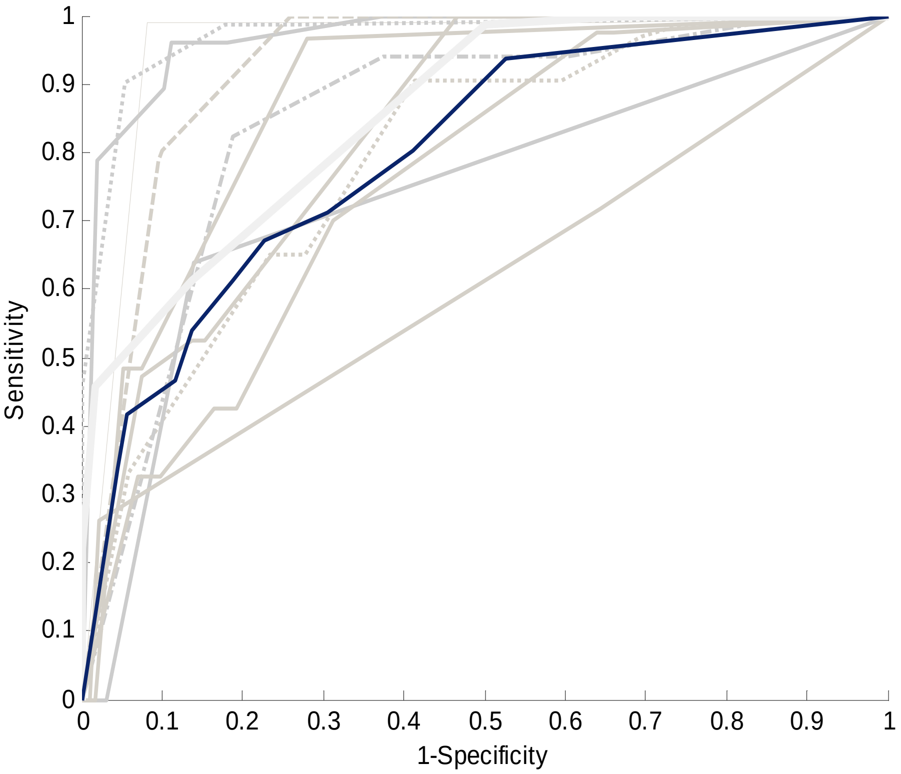


Figure 6. 6 ROC curves for 14 high grade patients' image classification. ROC curve of the classification performance for each patient was drawn with a thin line, and that of the average performance with a thick line.

6.4 DISCUSSION

In this study, we compared the diagnostic performance of three different imaging techniques; unpolarized white light, cross-polarized and parallel-polarized imaging, and reported the classification performance for the detection of cervical cancerous or precancerous lesions using cross polarized imaging which yielded the best image classification result among three imaging techniques. We believe that cross polarized imaging could provide the best performance since the specular reflection was removed through cross polarization. In our previous study, we found that unpolarized white light suffers from the limitation of not obtaining diagnostic information from areas obscured by specular reflection. Since the proposed MRF model in the previous study utilized diagnostic information from neighboring tissue regions, tissues surrounding the specular reflection areas could not be diagnosed properly. This limitation may reduce the diagnostic performance when unpolarized white light imaging is utilized in a MRF framework.

Additionally, we found that intensity contrast between squamous tissue and columnar tissue or the transformation zone was increased using cross polarized images. Cross polarized imaging provides information about deeper tissue structures, and, therefore, vascular patterns in the transformation zone could be measured more clearly. As a result, our automated segmentation algorithm could separate tissue types more successfully using cross polarized images than using unpolarized images. Since we trained our diagnosis algorithm differently depending on tissue type information, better tissue type separation directly resulted in better diagnostic performance.

This pilot study explored the diagnostic effectiveness of cross polarized imaging. A potential weakness of this study is the small sample size which make it difficult to generalize our preliminary findings. Moreover, due to the limited number of samples, we could not investigate the diagnostic value of abnormal vascular patterns such as punctation or hairpin vascular patterns. The vascular patterns are a prominent feature of neoplastic lesions [38]. We expect that the performance of our diagnostic algorithm will be increased by incorporating vascular pattern analysis algorithm. We will continue to investigate vascular patterns in larger scale studies in the future.

A major strength of the proposed approach is that it applied an optical technology in order to overcome a practical difficulty of the conventional screening methods.

Our automated approach allows a less trained provider to effectively identify the transformation zone and to quantitatively assess the margin of neoplastic lesions. This suggests that automated diagnostic tools may help provide a cost effective solution for screening in many developing countries. If the proposed approach can be further improved with incorporating a vascular feature analysis algorithm, this technique may assist see and treat procedures with a high diagnostic performance in low resource settings.

Chapter 7: Conclusion

7.1 SUMMARY OF RESULTS

This dissertation proposed a diagnostic image analysis framework using MDCs for early detection of cervical cancer. The proposed framework is intended to provide a low cost screening tool as a suitable alternative to conventional screening techniques, especially in developing countries. Based on the proposed framework, various studies were performed and the diagnostic performance of each study was quantitatively reported. For the quantitative diagnostic performance measure including sensitivity and specificity in the image classification context, we developed a gold-standard based approach.

A pilot study using white light reflectance imaging was performed and showed promising agreement between the image analysis results and histopathology with sensitivity of 79% and specificity of 88% from twenty nine patients. This is comparable to the performance of conventional colposcopy where the sensitivity and specificity have been reported to range from 87 – 96 % and 34 – 85 % [74-76].

In order to consider the information about tissue type, a MRF model-based approach was applied for the detection of cervical cancerous or precancerous lesions using multispectral digital colposcope images. The image analysis results suggest that the tissue type information is a diagnostically significant feature and that the features selected from reflectance and from fluorescence imaging show tissue-type selective diagnostic performance. A MRF model-based approach of this study left a concern that specular reflection might have reduced the diagnostic performance of our method.

In order to remove specualr reflection and improve the diagnostic performance in a MRF-based framework, polarized imaging techniques were applied and the results were compared to those using unpolarized imaging. The results showed that the classification performance for the detection of cervical cancerous or precancerous lesions using cross polarized imaging was the best among three imaging techniques including cross polarized, parallel polarized, and unploarized imaging.

7.2 FUTURE DIRECTIONS

The research presented in this dissertation indicates the potential of an automated diagnostic image analysis tool to detect cervical cancerous or precancerous lesions. A potential weakness of this study is relatively small number of sample sizes which makes it difficult to generalize our preliminary findings. Larger scale clinical studies would be needed in the future.

Our approach is limited to the detection of neoplastic lesions on the ectocervix only. In order to find lesions located on the endocervix, additional screening techniques will be required. Future studies will consider such additional screening techniques for lesions on the endocervix and incorporate them into the proposed framework.

Another practical problem in this study is the difficulty in maintaining proper focus with the instrument in the clinical settings primarily due to patients' movement during measurement. This may degrade the diagnostic performance of our approach since the performance of image analysis is highly dependent on image quality. One of the solutions would be to develop a system with auto focusing function, which can be designed in the further study.

By reducing interactions with clinical personnel, automated screening solutions may provide a more cost effective screening program for low-resource setting. This makes such solutions more suitable for clinical applications in developing countries. The proposed approach laid the solid ground for an automated diagnostic image analysis tool. Following the proposed framework, we plan to study on further automation of the proposed approach in the future.

References

[1] American Cancer Society. 2007. Cancer facts and figures 2007.

[2] World Health Organization. 2002. Cervical cancer screening in developing countries. Report of a WHO consultation.

[3] Agosti, J. M., S. J. Goldie. 2007. Introducing HPV vaccine in developing countries -key challenges and issues. *N Engl J Med.* 365: 1908-10.

[4] Parkin, D. M., F. Bray, J. Ferlay, and P. Pisani. 2005. Global cancer statistics. *CA Cancer J Clin.* 55: 74-108.

[5] Sodhani, P., S. Gupta, J. K. Sharma, A. Parashari, K. Halder, V. Singh, and A. Sehgal. 2006. Test characteristics of various screening modalities for cervical cancer: a feasibility study to develop an alternative strategy for resource-limited settings. *Cytopathology* 17 (6): 348-52.

[6] Parashari, A., V. Singh, A. Sehgal, L. Satyanarayana, P. Sodhani, M.M. Gupta. 2000. Low-cost technology for screening uterine cervical cancer. *Bulletin of the World Health Organization* 78 (8): 964-67.

[7] Goldie, S. J., G. Lynne, D. Jeremy, F. Goldhaber, G. T. Amparo, L. Carol, et al. 2005. Cost-Effectiveness of Cervical-Cancer Screening in Five Developing Countries. *NEJM* 353: 2158-68.

[8] Nene, B. M., S. Deshpande, K. Jayant, A. M. Budukh, P. S. Dale, D. A. Deshpande, et al. 1996. Early detection of cervical cancer by visual inspection: a population-based study in rural India. *Int. J. Cancer* 68: 770-3.

[9] Wesley, R., R. Sankaranarayanan, B. Mathew, B. Chandralekha, A. Aysha-Beegum, N. S. Amma, et al. 1997. Evaluation of visual inspection as a screening test for cervical cancer. *Br. J. Cancer* 75: 436-40.

[10] Basu, P., R. Sankaranarayanan, R. Mandal, C. Roy, P. Das, D. Choudhury, et al. 2002. Evaluation of downstaging in the detection of cervical neoplasia in Kolkata, India. *Int. J. Cancer* 100: 92-6.

[11] Germar, M.J. and M. Merialdi. 2006. Visual inspection with acetic acid as a cervical cancer screening tool for developing countries. Review for postgraduate

training course in reproductive health/chronic disease 2003, (September 14, 2006), http://www.gfmer.ch/.

[12] Sankaranarayanan, R., S. S. Shastri, P. Basu, C. Mahe, R. Mandal, G. Amin, et al. 2004. The role of low-level magnification in visual inspection with acetic acid for the early detection of cervical Neoplasia. *Cancer Detect Prev.* 28: 345-51.

[13] Denny, L., L. Kuhn, A. Pollack, and T. Wright. 2002. Direct visual inspection for cervical cancer screening: an analysis of factors influencing test performance. *Cancer* 94: 1699-707.

[14] Sankaranarayanan, R., R. Rajkumar, R. Theresa, P. O. Esmy, C. Mahe, K. R. Bagyalakshmi, S. Thara, L. Frappart, E. Lucas, R. Muwonge, S. Shanthakumari, D. Jeevan, T. M. Subbarao, D. M. Parkin, and J. Cherian. 2004. Initial results from a randomized trial of cervical visual screening in rural south India. *Int J Cancer* 109: 461-7.

[15] Jeronimo, J., O. Morales, J. Horna, J. Pariona, J. Manrique, J. Rubinos, and R. Takahashi. 2005. Visual inspection with acetic acid for cervical cancer screening outside of low-resource settings. *Rev Panam Salud Publica.* 17 (1): 1-5.

[16] Ramanujam, N., M. F. Mitchell, A. Mahadevan, S. Warren, S. Thomsen, E. Silva, et al. 1994. In vivo diagnosis of cervical entraepithelial neoplasia using 337-nm-excited laser-induced fluorescence. *Proc. Natl. Acad. Sci. USA* 91: 10193-7.

[17] Ramanujam, N., M. F. Mitchell, A. Mahadevan, S. Thomsen, A. Malpica, T. Wright, et al. 1996. Spectroscopic diagnosis of cervical intraepithelial neoplasia (CIN) in vivo using laser-induced fluorescence spectra at multiple excitation wavelengths. *Lasers Surg. Med.* 19: 63-74.

[18] Ramanujam, N., M. F. Mitchell , A. Mahadevan-Jansen, A. Thomsen, G. Staerkel, A. Malpica, et al. 1996. Cervical precancer detection using a multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths. *Photochem. Photobiol.* 64: 720-35.

[19] Brookner, C., U. Utzinger , G. Staerkel , R. Richards-Kortum , and M. F. Mitchell. 1999. Cervical fluorescence of normal women. *Lasers Surg. Med.* 24: 29-37.

[20] Chang, S., M. Dawood ,G. Staerkel, U. Utzinger, R. Richards-Kortum, and M. F. Follen. 2002. Fluorescence spectroscopy for cervical pre-cancer detection: Is there variance across the menstrual cycle? *J. Biomed. Opt.* 7: 595-602.

[21] Drezek, R., C. Brookner, I. Pavlova , I. Boiko, A. Malpica, R. Lotan, et al. 2001. Autofluorescence microscopy of fresh cervical tissue sections reveals alterations in tissue biochemistry with dysplasia. *Photochem. Photobiol.* 73: 636-41.

[22] Ferris, D. G., R. A. Lawhead, E. D. Dickman, N. Holtzapple, J. A. Miller, S. Grogan, S. Bambot, A. Agrawal, and M. L. Faupel. 2001. Multimodal hyperspectral imaging for the noninvasive diagnosis of cervical neoplasia. *Journal of Lower Genital Tract Disease* 5(2): 65-72.

[23] Huh, W. K., R. M. Cestero, F. A. Garcia, M. A. Gold, R. M. Guido, K. McIntyre-Seltman, D. M. Harper, L. Burke, S. T. Sum, R. F. Flewelling, and R. D. Alvarez. 2004. Optical detection of high-grade cervical neoplasia in vivo: Results of a 604 patient study. *Am J Obstet Gynecol* 190: 1249-57.

[24] Ries, L. A. G., M. P. Eisner, C. L. Kosary, et al., eds. 2002. SEER cancer statistics review, 1973-1999. *Bethesda, Maryland: National Cancer Institute,* http://seer.cancer.gov/csr/1973_1999.

[25] Alvarez, R. D., T. C. Wright, and Optical Detection Group. 2007. Effective cervical Neoplasia detection with a novel optical detection system: A randomized trial. *Gynecol Oncol* 104: 281-9.

[26] Crisp, W. E., B. L. Craine, and E. A. Craine. 1990. The computerized digital imaging colposcope: future directions. *Am. J. Obstet. Gynecol.* 162: 1491-7.

[27] Craine, B. L. and E. R. Craine. 1993. Digital imaging colposcopy: basic concepts and applications. *Obstet. Gynecol.* 82: 869-73.

[28] Shafi, M. I., J. A. Dunn, R. Chenoy, E. J. Buxton, C. Williams, and D. M. Luesley. 1994. Digital imaging colposcopy, image analysis and quantification of the colposcopic image. *Br. J. Obstet. Gynaecol.* 101: 234-8.

[29] Cristoforoni, P. M., D. Gerbaldo, A. Perino, R. Piccoli, F. J. Montz, and G. L. Capitanio. 1995. Computerized colposcopy: results of a pilot study and analysis of its clinical relevance. *Obstet. Gynecol.* 85: 1011-6.

[30] Pogue, B. W., H. B. Kaufman, A. Zelenchuk, W. Harper, G. C. Burke, E. E. Burke, et al. 2001. Analysis of acetic acid-induced whitening of high-grade squamous intraepithelial lesions. *J. Biomed. Optics* 6: 397-403.

[31] Li, W., V. Van Raad, J. Gu, U. Hansson, J. Hakansson, H. Lange, and D. Ferris. 2005. Computer-aided Diagnosis (CAD) for cervical cancer screening and diagnosis: a new system design in medical image processing. *Lecture Notes in Computer Science, CVBIA 2005*: 240-50.

[32] Parker, M. F. 2005. Emerging technology in cervical cancer screening: spectroscopy. *Clinic. Obstet & Gynecol.* 48: 209-17.

[33] Balas, C. 2001. A novel optical imaging method for the early detection, quantitative grading, and mapping of cancerous and precancerous lesions of cervix. *IEEE T. Biomed. Eng.* 48: 96-104.

[34] Gordon, S., G. Zimmerman, and H. Greenspan. Image Segmentation of Uterine Cervix Images for Indexing in PACs. 2004. *in Proceedings of IEEE 17th Symposium on Computer-based Medical Systems*.

[35] Gordon, S., G. Zimmerman, R. Long, S. Antani, J. Jeronimo, and H. Greenspan. 2006. Content analysis of uterine cervix images: initial steps towards content based indexing and retrieval of cervigrams. *Medical Imaging 2006: Image Processing.* Edited by J. M. Reinhardt and J. P. Pluim, in Proc. SPIE 6144: 1549-56.

[36] Van Raad, V., Z. Xue, and H. Lange. 2006. Lesion margin analysis for automated classification of cervical cancer lesions. *Medical Imaging 2006: Image Processing.* Edited by J. M. Reinhardt and J. P. Pluim, in Proc. SPIE 6144: 1647-59.

[37] Ji, Q., J. Engel, and E. Craine. 2000. Texture Analysis for Classification of Cervix Lesions. *IEEE Trans. Med. Imaging* 19: 1144-9.

[38] Li, W. and A. Poisson. 2006. Detection and characterization of abnormal vascular patterns in automated cervical image analysis. *Lecture Notes in Computer Science: Advances in Visual Computing* 4292: 627-636.

[39] Benavides, J. M., S. Chang, S. Park, R. Richards-Kortum, N. Mackinnon, C. MacAulay, et al. 2003. Multispectral digital colposcope for in vivo detection of cervical cancer. *Optics Express* 11: 1223-36.

[40] Li, W., M. Soto-Thompson, and U. Gustafsson. 2006. A new image calibration system in digital colposcopy. *Optics Express* 14 (26): 12887-901.

[41] Elson, D. A., R. R. Riley, A. Lacey, G. Thordarson, F. J. Talamantes, and J. M. Arbeit. 2000. Sensitivity of the Cervical Transformation Zone to Estrogen-induced Squamous Carcinogenesis. *Cancer Research* 60: 1267–75.

[42] Zenklusen, J. C., S. L. Stockman, S. M. Fischer, C. J. Conti, and I. B. Gimenez-Conti. 1994. Transforming growth factor-beta 1 expression in Syrian hamster cheek pouch carcinogenesis. *Mol. Carcinog.* 9: 10-16.

[43] Gimenez-Conti, I. B., D. M. Shin, A. B. Bianchi, D. R. Roop, W. K. Hong, C. J. Conti, and T. J. Slaga. 1990. Changes in keratin expression during 7,12-dimethylbenz[a]anthracene-induced hamster cheek pouch carcinogenesis. *Cancer Res.* 50: 4441-5.

[44] Coghlan, L., U. Utzinger, R. Richards-Kortum, C. Brookner, A. Zuluaga, I. Gimenez-Conti, M. Follen-Mitchell. 2001. Fluorescence spectroscopy of epithelial tissue throughtout the dysplasia-carcinoma sequence in an animal model: spectroscopic changes precede morphologic changes. *Lasers Surg. Med.* 29: 1-10.

[45] Coghlan, L., U. Utzinger, R. Drezek, D. Heintzelman, A. Zuluaga, C. Brookner, R. Richards-Kortum. 2000. Optimal fluorescence excitation wavelengths for detection of squamous intra-epitthelial neoplasia:results from an animal model. *Opt. Express* 7: 436-45.

[46] Sevick-Muraca, E., and R. Richards-Kortum. 1996. Quantitative optical spectroscopy for tissue diagnosis. *Ann Rev Phys. Chem.* 47: 555-606.

[47] Wagnieres, G. A., W. M. Star, and B. C. Wilson. 1998. In vivo fluorescence spectroscopy and imaging for oncological applications. *Photochem. Photobiol.* 68: 603-32.

[48] Stepp, H., R. Sroka, and R. Baumgartner. 1998. Fluorescence endoscopy of gastrointestinal diseases: basic principles, techniques, and clinical experience. *Endoscopy* 30: 379-86.

[49] Cantor, S. B., M. Follen-Mitchell, G. Tortolero-Luna, C. Bratka, D. Bodurka, and R. Richards-Kortum. 1998. Cost-effectiveness analysis of diagnosis and management of cervical squamous intraepithelial lesions. *Obstet. Gynecol.* 91: 270-77.

[50] Gillenwater, A. R., R. Jacob, R. Ganeshappa, B. Kemp, A. K. El-Naggar, J. L. Palmer, G. Clayman, M. Follen-Mitchell, and R. Richards-Kortum. 1998. Noninvasive diagnosis of oral neoplasia based on fluorescence spectroscopy and native tissue autofluorescence. *Arch. Otolaryngol. Head Neck Surg.* 124: 1251-8.

[51] Ramanujam, N., M. Follen-Mitchell, A. Mahadevan-Jansen, S. Thomsen, G. Staerkel, A. Malpica, T. Wright, N. Atkinson and R. Richards-Kortum. 1996. Cervical precancer detection using multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths. *Photochem. Photobiol.* 64: 720-735.

[52] Agrawal, A., U. Utzinger, C. Brookner, C. Pitris, M. Follen-Mitchell, and R. Richards-Kortum. 1999. Fluorescence spectroscopy of the cervix: influence of acetic acid, cervical mucus, and vaginal medications. *Lasers Surg. Med.* 25: 237-49.

[53] Onizawa, K., H. Saginoya, Y. Furuya and H. Yoshida. 1996. Fluorescence photography as a diagnostic method for oral cancer. *Cancer Lett.* 108: 61-6.

[54] Dhingra, J. K., X. Zahng, K. McMillan, S. Kabani, R. Manoharan, I. Itzkan, M. S. Feld, S. M. Sharpshay. 1998. Diagnosis of head and neck precancerous lesions in an animal model using fluorescence spectroscopy. *Laryngoscope* 108: 471-5.

[55] Wang, C. Y., T. Tsai, H. C. Chen, S. C. Chang, C. T. Chen, and C. P. Chiang. 2003. Autofluorescence spectroscopy for in vivo diagnosis of DMBA-induced hamster buccal pouch pre-cancers and cancers. *J. Oral Pathol. Med.* 32: 18–24.

[56] Vengadesan, N., P. Aruna, and S. Ganesan. 1998. Characterization of native fluorescence from DMBA-treated hamster cheek pouch buccal mucosa for measuring tissue transformation. *Br. J. Cancer* 77: 391-5.

[57] Chen, C. T., H. K. Chiang, S. N. Chow, C. Y. Wang, Y. S. Lee, J. C. Tsai, and C. P. Chiang. 1998. Autofluorescence in normal and malignant human oral tissues and in DMBA-induced hamster buccal pouch carcinogenesis. *J. Oral Pathol. Med.* 27: 470-4.

[58] Heintzelman, D. L., U. Utzinger, et al. 2000. Optimal excitation wavelengths for in vivo detection of oral neoplasia using fluorescence spectroscopy. *Photochem. Photobiol.* 72: 103-13.

[59] Muller, M. G., T. A. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, N. Laver, Z. Wang, C. W. Boone, R. R. Dasari, S. M. Shapshay, and M. S. Feld. 2003. Spectroscopic detection and evaluation of morphologic and biochemical changes in early human oral carcinoma. *Cancer* 97: 1681-92.

[60] Bhaskaran, V. and K. Konstantinides. 1999. *Image and video compression standards algorithms and architectures.* USA: Kluwer Academic Publishers.

[61] Chang, S. K., M. Follen, A. Malpica, U. Utzinger, S. Gtaerkel, D. Cox, N. Atkinson, C. MacAulay, and R. Richards-Kortum. 2002. Optimal excitation wavelengths for discrimination of cervical neoplasia. *IEEE Trans. Biomed. Eng.* 49: 1102-11.

[62] Fei, B. W., Z. H. Lee, D. T. Boll, J. L. Duerk, D. B. Sodee, J. S. Lewin, and D. L. Wilson. 2004. Registration and fusion of SPECT, high-resolution MRI, and interventional MRI for thermal ablation of prostate cancer. *IEEE Trans. Nucl, Science* 51 (1): 177-83.

[63] Maes, F., A. Collignon, D. Vandermeulen, G. Marchal, and P. Seutens. 1997. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging* 16: 187-98.

[64] Hsu, R., M. Abdel-Mottaleb, and A. K. Jain. 2002. Face detection in color images. *IEEE Trans. Patt. Anal. Mach. Intell.* 24: 696-706.

[65] Richard, O. D., E. H. Peter, and G. S. David. 2001. *Pattern Classification.* John Wiley & Sons, Inc.

[66] Bera, A. and C. Jarque. 1981. Efficient tests for normality, heteroskedasticity and serial independence of regression residuals: Monte Carlo evidence. *Econo. Letter* 7: 313-18.

[67] Inaguma, M. and K. Hashimoto. 1999. Porphyrin-like fluorescence in oral cancer. *Cancer* 86: 2201-11.

[68] Koss, L. G.. 1989. The Papanicolaou test for cervical cancer detection. A triumph and a tragedy. *JAMA* 261: 737-43.

[69] Bomfim-Hyppolito, S., E. S. Franco, R. G. Franco, C. M. de Albuquerque, and G. C. Nunes. 2006. Cervicography as an adjunctive test to visual inspection with acetic acid in cervical cancer detection screening. *Intl. J. Gynecol. & Obstet* 92: 58-63.

[70] Park, S., T. Collier, J. Aaron, M. Markey, K. Sokolov, R. Richards-Kortum, et al. 2005. Multispectral digital microscopy for In Vivo monitoring of oral neoplasia in the hamster cheek pouch model of carcinogenesis. *Optics Express* 13: 749-62.

[71] Boonstra, H., J. W. Oosterhuis, A. M. Oosterhuis, and G. J. Fleuren. 1983. Cervical tissue shrinkage by formaldehyde fixation, paraffin wax embedding, section cutting and mounting Virchows Arch. *A Pathol. Anat. Histopatho.* 402: 195-201.

[72] Kohavi, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI*: 1137-43. San Mateo, CA.

[73] Aubin, J. P. 1998. *Optima and Equilibria: An introduction to nonlinear analysis*, 2nd ed. Berlin: Springer-Verlag.

[74] Follen, M. M., D. Schoteenfeld, G. Tortolero-Luna, S. B. Cantor, and R. Richard-Kortum. 1998. Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis. *Obstetrics & Gynecology* 91: 626-31.

[75] Benedet, J. L., G. H. Anderson, J. P. Mastic, and D. M. Mille. 1991. A quality-control program for colposcopic practice. *Obstet & Gynecol.* 78: 872-5.

[76] Seshadri, L., P. Jairaj, H. Krishnaswami. 1990. Colposcopy in the diagnosis of cervical Neoplasia. *Indian Journal of Cancer* 27: 180-6.

[77] Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston: Kluwer Academic Publishers.

[78] Shannon, C. E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal* 27: 379-423, 623-56.

[79] Davies, D. L. and D. W. Bouldin. 1979. A cluster separation measure. *IEEE Trans. Pattern Anal. Machine Intell.* 1: 224-7.

[80] Geman, S. and D. Geman. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6:721-41.

[81] Li, S. Z. 2001. *Markov random field modeling in image analysis*. Secaucus: Springer-Verlag New York, Inc.

Vita

Sun Young Park was born in Seoul, Korea, on May 17, 1972, the daughter of Jong Bae Park and Young Sook Kim. She graduated from Ewha girls' high school in 1991, and entered the department of Computer Science in Ewha Womans' University. After she got her Bachelor's degree in 1996, she continued to pursue a Master of Science degree in Computer Science at Ewha Womans' University. She researched the tumor detection algorithm using MRI data under the supervision of Dr. Myung Hee Kim. From June 1995, she has been an exchange researcher in Fraunhofer, Darmstadt, Germany for six months. After she received her Master of Science degree in February 1996, she was employed at Harris, Seoul, Korea as a software engineer for integrating the InChon international airport system. In 2003, she entered the Ph.D. program in Biomedical Engineering at the University of Texas at Austin, Austin, TX under the supervision of Dr. Rebecca Richards-Kortum.

Permanent address: Samhwa billa #410, 391-220, Daebang-dong, Dongjak-gu,
Seoul, Korea 156-020

This dissertation was typed by the author.